


```

sequence.
AC006262 AC006262 GI:15718550
VERSION AC006262.2
KEYWORDS HTG.
SOURCE
ORGANISM Homo sapiens
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 106320)
REFERENCE
AUTHORS Lamerdin,J.E., McCreedy,P.M., Skowronski,E., Viswanathan,V.,
Burkhardt-Schultz,K., Gordon,L., Dias,J., Ramirez,M., Stillwagen,S.,
Phan,H., Velasco,N., Do,L., Regala,M., Terry,A., Garnez,J.,
Dangnan,L., Erlar,A., Christensen,M., Georgescu,A., Arila,J.,
Liu,S., Atlix,C., Andreise,T., Frankheim,M., Amico-Keller,G.,
Krommiller,B., Duarte,S., Lucas,S., Bruce,R., Thomas,P., Quan,G.,
Kobayashi,A., Olsen,A.S. and Carrano,A.V.
Sequence analysis of a 1.9 Mb region in 19q13.2 between APOE and
D19S412
TITLE
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 107573)
AUTHORS Lamerdin,J.E.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1998) Joint Genome Institute, Lawrence Livermore
National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
3 (bases 1 to 107573)
REFERENCE
AUTHORS Lamerdin,J.E.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2001) Joint Genome Institute, Lawrence Livermore
National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
On Sep 21, 2001 this sequence version replaced gi:4079612.
Map and sequence oriented from centromere to q telomere. BC264576
overlaps BC282485 (CFC-344H19, AC007785) to the left from bases 1
to 20,296 of this accession, and overlaps BC82621 (CTB-139A18,
AC007193) to the right from bases 106,317 to 107,573 of this
accession. This sequence does not represent the entire insert of
BC264576. Additional chromosome 19 map and sequence information may
be obtained at: http://www.bio.lnl.gov/dbp/genome/genome.html.
Location/Qualifiers
1. 107573
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2 between APOE and D19S412"
/clone="CFC-297N14"
/cell_line="978SK"
/note="LIML clone name: BC264576 BAC library obtained
from Research Genetics."
/feature
misc_feature
354..567
/note="DDB similarity to AA456254 zx99d06.r1 Soares
NHMPUS1 Homo sapiens CDNA clone 811883 5'; (3..215); 100%
identity."
repeat_region
complement(669..800)
/rpt_family="(GA)n"
repeat_region
complement(829..953)
/rpt_family="(GGGA)n"
repeat_region
complement(1035..2017)
/rpt_family="LIMB7"
repeat_region
complement(2014..2202)
/rpt_family="LIMB7"
repeat_region
complement(2206..2495)
/rpt_family="AluJo"
repeat_region
complement(2497..2728)
/rpt_family="LIMB7"
repeat_region
complement(2730..3133)
/rpt_family="LIMB7"
repeat_region
complement(3158..3447)
/rpt_family="AluJo"
repeat_region
complement(3448..3803)
/rpt_family="LIMB7"
repeat_region
complement(3805..4108)
/rpt_family="AluJo"

```

```

repeat_region complement(4110..4801)
/rpt_family="LIMB7"
repeat_region complement(4826..5069)
/rpt_family="LIMB7"
repeat_region complement(5338..5369)
/rpt_family="(7A)n"
repeat_region complement(5370..5708)
/rpt_family="LIMB7"
repeat_region complement(5974..6277)
/rpt_family="AluY"
repeat_region complement(7059..7080)
/rpt_family="AT-rich"
repeat_region complement(7150..7485)
/rpt_family="L1"
repeat_region complement(7527..7646)
/rpt_family="FLAM-C"
repeat_region complement(7661..7916)
/rpt_family="L1"
repeat_region 7938..9725
/rpt_family="L1"
repeat_region complement(9736..10038)
/rpt_family="AluSg"
misc_feature complement(10866..10964)
/note="predicted exon, program: graal2exons_human_1.3,
frame: 1, quality: excellent, score: 75.000"
repeat_region 11326..11517
/rpt_family="AluJo"
misc_feature 11945..11998
/note="DDB similarity to AA456254 zx99d06.r1 Soares
NHMPUS1 Homo sapiens CDNA clone 811883 5'; (210..269);
100% identity."
misc_feature 12144..12411
/note="DDB similarity to overlapping ESTs:
(12144..12398) AA456254 zx99d06.r1 Soares NHMPUS1 Homo
sapiens CDNA clone 811883 5'; (270..523); 98%
identity.-(12182..12411) AT081371 ox76g10.x1
Soares.NHMPUS1 Homo sapiens CDNA clone IMAGE:1662306 3';
(454..225); 100% identity.-(12173..12411) AA454642
zx99d06.r1 Soares NHMPUS1 Homo sapiens CDNA clone 811883
3'; (455..217); 100% identity.-(12144..12411) AA760247
vv74e06.r1 Stragene mouse skin (#937313) Mus musculus
CDNA clone 1228162 5'; (119..380); 69% identity."
misc_feature complement(12536..12614)
/note="predicted exon, program: graal2exons_human_1.3,
frame: 2, quality: excellent, score: 90.000"
misc_feature 12600..12818
/note="DDB similarity to overlapping ESTs:
(12603..12818) AA45642 zx99d06.r1 Soares NHMPUS1 Homo
sapiens CDNA clone 811883 3'; (216..1); 99%
identity.-(12600..12823) AT081371 ox76g10.x1
Soares.NHMPUS1 Homo sapiens CDNA clone IMAGE:1662306 3';
(224..1); 98% identity.-(12600..12656) AA760247 vv74e06.r1
Stragene mouse skin (#937313) Mus musculus CDNA clone
1228162 5'; (381..432); 76% identity."
repeat_region complement(12967..16062)
/rpt_family="L1P5"
misc_feature complement(14029..14445)
/note="predicted exon, program: graal2exons_human_1.3,
frame: 0, quality: good, score: 58.000"
repeat_region complement(17695..18620)
/rpt_family="MER11B"
repeat_region complement(18619..18767)
/rpt_family="MER11A"
repeat_region 19151..19519
/rpt_family="MER11B"
repeat_region complement(19592..19754)
/rpt_family="AluJo"
repeat_region complement(19767..19870)
/rpt_family="MER65A"
repeat_region complement(20427..20481)
/rpt_family="(CA)n"
repeat_region 20546..20751
/rpt_family="AluJo"

```

```

repeat_region      21128..21397
                    /rpt_family="Alusx"
repeat_region      21768..22798
                    /rpt_family="MER11A"
misc_feature       complement(23090..23194)
                    /note="predicted exon, program: grail2exons_human_1.3,
                    frame: 2, quality: good, score: 50.000"
repeat_region      complement(24023..24141)
                    /rpt_family="FLAM.A"
repeat_region      complement(24258..24429)
                    /rpt_family="AluJb"
repeat_region      complement(24910..25309)
                    /rpt_family="MSTB"
repeat_region      25329..25464
                    /rpt_family="HERVL"
repeat_region      25732..25881
                    /rpt_family="MER4.5"
repeat_region      complement(26586..26613)
                    /rpt_family="AT-rich"
repeat_region      complement(26677..26941)
                    /rpt_family="MIR"
repeat_region      complement(27078..27369)
                    /rpt_family="Alusx"
repeat_region      complement(27374..27682)
                    /rpt_family="Alusg"
repeat_region      27743..28177
                    /rpt_family="L1M1/2"
repeat_region      28646..28929
                    /rpt_family="Alusx"
repeat_region      28987..29070
                    /rpt_family="Alusx"
repeat_region      29085..29341
                    /rpt_family="(TA)n"
misc_feature       /note="predicted exon, program: grail2exons_human_1.3,
                    frame: 2, quality: good, score: 50.000"
repeat_region      29332..29526
                    /rpt_family="MER4D"
repeat_region      29527..29826
                    /rpt_family="Alusg"
repeat_region      29827..29870
                    /rpt_family="MER4D"
repeat_region      29870..30491
                    /rpt_family="MER4D"
repeat_region      30523..30611
                    /rpt_family="(GA)n"
repeat_region      31175..31465
                    /rpt_family="Alusg"
repeat_region      complement(31707..32005)
                    /rpt_family="Alusg"
repeat_region      complement(32050..32614)
                    /rpt_family="MLT2D"
repeat_region      32619..32691
                    /rpt_family="LTR8"

Query Match      41.7%; Score 318; DB 9; Length 107573;
Best Local Similarity 99.5%; Pred. No. 1.5e-175;
Matches 418; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 584 TGCTGATGGGGGAGGAGGAGCTGTAAACCTCTGTATGACCCCTATGGCCAAATCAACC 643
      |||||||
DB 82388 TGCTGATGGGGGAGGAGGAGCTGTAAACCTCTGTATGACCCCTATGGCCAAATCAACC 82647
QY 644 CGGACACACCCCAAGGCTGGTGGGAACCTTACCTTGTGAGATTTTCCATCATC 703
      |||||||
DB 82648 CGGACACACCCCAAGGCTGGTGGGAACCTTACCTTGTGAGATTTTCCATCATC 82707
QY 704 TCAAGTTCTCTTATCCAGAGACCAACAGATCAATTAATTAATGACTTTATA 763
      |||||||
DB 82708 TCAAGTTCTCTTATCCAGAGACCAACAGATCAATTAATTAATGACTTTATA 82767

RESULT 4
AC007193/c 146180 bp DNA linear PRI 17-JUN-1999
LOCUS Homo sapiens chromosome 19, BAC 82621 (CIT-B-139a18), complete
DEFINITION sequence.
ACCESSION AC007193
VERSION AC007193.1 GI:4558635
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 146180)
AUTHORS Lamerdin,J.E., McCreedy,P.M., Skowronski,E., Visevathian,V.,
Burkhardt-Schultz,K.J., Gordon,L., Dias,J., Ramirez,M.,
Stilwagen,S., Phan,H., Velasco,N., Do,L., Regala,W., Terry,A.,
Garnes,J., Danganan,L., Erler,A., Christensen,M., Georgescu,A.,
Avila,J., Liu,S., Attix,C., Andrade,S., Lucas,S., Bruce,R.,
Amico-Keller,G., Coefield,J., Duarte,S., Sanders,C., Ow,D.,
Thomas,P., Quan,G., Krommiller,B., Arellano,A., Olsen,A.S., and Carrano,A.V.,
Nolan,M., Truong,S., Kobayashi,A., Olsen,A.S., and Carrano,A.V.,
Sequence analysis of a 1.9 Mb region in 19q13.2 between APOE and
D19S412
TITLE Unpublished
JOURNAL 2 (bases 1 to 146180)
REFERENCE Lamerdin,J.E.
AUTHORS Direct Submission
TITLE Submitted (02-Apr-1999) Joint Genome Institute, Lawrence Livermore
NATIONAL LABORATORY, 7000 East Ave., Livermore, CA 94551, USA
COMMENT (CIT-B-139a18) overlaps BAC 264576 (CIT-B-297n14) to the left from
bases 1 to 27,487 of this accession, and separated from BAC 89981
(CIT-B-158d10) on the right by a gap of approximately 17 to 23 kb.
Additional chromosome 19 map and sequence information may be
obtained at: http://www.bio.lnl.gov/dbp/genome/genome.html.

FEATURES
Source
1..146180
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2 between APOE and D19S412"
/clone="BC82621"
/cell_line="987SK"
/cell_type="fibroblast"
/note="LNL clone name: BC82621 BAC library obtained from
Research Genetics."
1..101
/rpt_family="L1"
repeat_region 511..677
/rpt_family="L1"
repeat_region 1018..1158
/rpt_family="L1"
repeat_region 1159..1453
/rpt_family="Alusx"
repeat_region 1459..1529
/rpt_family="L1"
repeat_region complement(1857..1974)
/rpt_family="FLAM.A"
repeat_region 2115..2375
/rpt_family="L1M4"

```

```
repeat_region 2478..2549 /rpt_family="LINE2"  
repeat_region complement(2629..2921)  
repeat_region /rpt_family="AluJo"  
repeat_region 3355..3655 /rpt_family="AluSg"  
repeat_region 3673..3939 /rpt_family="L1"  
repeat_region complement(4005..4264)  
repeat_region /rpt_family="AluSx"  
repeat_region complement(4269..4568)  
repeat_region /rpt_family="AluY"  
misc_feature 5064..5122 /note="predicted exon, program: graal2exons_human_1.3,  
frame: 2, quality: excellent, score: 81.000"  
repeat_region /rpt_family="AluY"  
misc_feature 5626..5686 /note="predicted exon, program: graal2exons_human_1.3,  
frame: 0, quality: good, score: 52.000"  
repeat_region 7398..7444 /rpt_family="Alu"  
repeat_region 7480..7595 /rpt_family="TRA)n"  
repeat_region 7596..7812 /rpt_family="AluJb"  
repeat_region 7900..8191 /rpt_family="AluSx"  
repeat_region complement(8575..8597)  
repeat_region /rpt_family="AT-rich"  
repeat_region complement(8598..8878)  
repeat_region /rpt_family="AluSx"  
repeat_region 9477..9766 /rpt_family="AluSx"  
repeat_region 9997..10258 /rpt_family="AluSg"  
repeat_region 10273..10347 /rpt_family="FLAM"  
repeat_region 10691..10818 /rpt_family="AluJb"  
repeat_region complement(10999..11283)  
repeat_region /rpt_family="AluJo"  
repeat_region complement(11302..11517)  
repeat_region /rpt_family="L1MC/D"  
repeat_region complement(11536..11661)  
repeat_region /rpt_family="FLAM_C"  
repeat_region complement(11667..11798)  
repeat_region /rpt_family="L1MC/D"  
repeat_region complement(11833..12138)  
repeat_region /rpt_family="AluSg"  
repeat_region complement(12158..12478)  
repeat_region /rpt_family="L1MC/D"  
repeat_region 12479..12642 /rpt_family="FRAM"  
repeat_region complement(12655..12795)  
repeat_region /rpt_family="L1MC/D"  
repeat_region complement(12806..13106)  
repeat_region /rpt_family="AluSx"  
repeat_region complement(13356..13481)  
repeat_region /rpt_family="(GAA)n"  
repeat_region complement(13482..13694)  
repeat_region /rpt_family="AluJb"  
repeat_region complement(13709..14037)  
repeat_region /rpt_family="L1"  
repeat_region complement(14009..14105)  
repeat_region /rpt_family="L1MC/D"  
repeat_region complement(14115..14415)  
repeat_region /rpt_family="AluY"  
repeat_region complement(14417..14849)  
repeat_region /rpt_family="L1MC/D"  
repeat_region complement(14859..15153)  
repeat_region /rpt_family="AluSx"  
repeat_region 15226..15492
```

```
misc_feature /rpt_family="AluJo"  
complement(16257..16340)  
/note="predicted exon, program: graal2exons_human_1.3,  
frame: 0, quality: good, score: 63.000"  
repeat_region 16362..16666 /rpt_family="AluSg"  
repeat_region complement(16691..17060)  
repeat_region /rpt_family="THEIC"  
repeat_region 17482..17781 /rpt_family="AluJo"  
repeat_region 17782..17870 /rpt_family="AluSg"  
repeat_region /rpt_family="(GAA)n"  
repeat_region 17968..18273 /rpt_family="AluSg"  
repeat_region complement(18401..18817)  
repeat_region /rpt_family="L1R7"  
repeat_region complement(18818..19193)  
repeat_region /rpt_family="HERVH"  
repeat_region complement(19194..19412)  
repeat_region /rpt_family="HERVH"  
repeat_region complement(19414..20028)  
repeat_region /rpt_family="HERVH"  
repeat_region complement(20031..23647)  
repeat_region /rpt_family="HERVH"  
repeat_region complement(23651..24066)  
repeat_region /rpt_family="L1R7"  
repeat_region complement(24403..24714)  
repeat_region /rpt_family="MER74"  
repeat_region 24785..25080 /rpt_family="AluSx"  
repeat_region 25161..25207 /rpt_family="L1MD3"  
repeat_region 25224..25509 /rpt_family="AluSg/X"  
repeat_region 25605..25901 /rpt_family="AluSg/X"  
repeat_region 25902..26069 /rpt_family="AluSg/X"  
repeat_region complement(26224..26256)  
repeat_region /rpt_family="(TRA)n"  
repeat_region complement(26257..26545)  
repeat_region /rpt_family="AluJo"  
repeat_region 26573..26876 /rpt_family="AluSg"  
repeat_region complement(27125..27422)  
repeat_region /rpt_family="AluSx"  
repeat_region 27866..28173 /rpt_family="AluSx"  
repeat_region 28181..28279 /rpt_family="AluSg/FLAM"  
repeat_region complement(28359..28522)  
repeat_region /rpt_family="AluSg"  
repeat_region complement(28526..28705)  
repeat_region /rpt_family="AluSg"  
repeat_region complement(28733..28872)  
repeat_region /rpt_family="AluJb"  
repeat_region 28874..29081 /rpt_family="AluSg"  
repeat_region 29463..29583 /note="predicted exon, program: graal2exons_human_1.3,  
frame: 2, quality: good, score: 71.000"  
repeat_region complement(29585..29888)  
repeat_region /rpt_family="AluSg"  
repeat_region 30336..30634 /rpt_family="AluSg"  
repeat_region 30646..30779 /rpt_family="AluSg"  
repeat_region complement(30830..31040)  
repeat_region /rpt_family="MIR"
```

Query Match 3.88; Score 29; DB 9; Length 146180;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	74	CAGGCTCCTCTGCACACGGAGCCCGAC	102
D6	6043	CAGGCTCCTCTGTCTCACGAGGCCCGAC	6015
RESULT 5			
AXI61419			
LOCUS	AXI61419	48 bp	DNA
DEFINITION	Sequence 21 from Patent WO0140465.	Linear	PAT 22-JUN-2001
ACCESION	AXI61419		
VERSION	AXI61419.1	GI:14545090	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match	2.8%; Score 21; DB 6; Length 48;		
Best Local Similarity	100.0%; Pred. No. 1.8;		
Matches	21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	104	GCCCCCATGACTCTTACT	124
D6	28	GCCCCCATGACTCTTACT	48
RESULT 6			
AC097591/c			
LOCUS	AC097591	86606 bp	DNA
DEFINITION	Rattus norvegicus clone CHZ30-170A15. *** SEQUENCING IN PROGRESS	Linear	HTG 20-DEC-2001
ACCESION	AC097591		
VERSION	AC097591.3	GI:17970407	
KEYWORDS	HTG: HTGS_PHASE1.		
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus		
	1 (bases 1 to 86606)		
	Murphy,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C.,		
	Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,D.,		
	Benton,J., Blmage,K., Blankenburg,K., Bonnin,D., Bouck,J.,		
	Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,		
	Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,		
	Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,		
	Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,		
	Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,		
	Day-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,		
	Denn,A.L., Ding,Y., Dinh,H.H., Douhaite,K.J., Draper,H.,		
	Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C.,		
	Elhaj,C., Escotto,M., Falls,T., Ferriguto,D., Flagg,N., Ford,J.,		
	Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,		
	Garta,R., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S.,		
	Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A.,		
	Hernandez,J., Hernandez,O., Hodgson,B., Hognes,M., Hollway,C.,		
	Hollins,B., Homsi,F., Howard,S., Huber,J., Huylk.S., Hume,J.,		
	Jackson,L.E., Jacobson,B., Jle,X., Johnson,R., Jolivert,S.,		
	Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,T.,		

Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, M., Louie, G., Louie, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M. P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokwu, S., Oguh, M., Okunnu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Pimus, E., Pu, L. L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolle, M., Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shooshani, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Swatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmali, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and GIBBS, R.

Direct Submission
 Submitted (20-OCT-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

On Dec 20, 2001 this sequence version replaced gi:17062373.

```

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: GGIU
Center clone name: CH230-170A15
-----
Summary Statistics

Assembly program: Phrap; version 0.990329First call to
findPhrapLast

Consensus quality: 75643 bases at least Q40
Consensus quality: 80572 bases at least Q30
Consensus quality: 85107 bases at least Q20
Estimated insert size: 68482; sum-of-ctrls estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 0.8x in Q20 bases; sum-of-ctrls estimation
-----

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 43 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 3236: contig of 3236 bp in length
* 3237
* 3336: gap of unknown length
* 3337
* 8095: contig of 4759 bp in length
* 8096
* 8195: gap of unknown length
* 8196
* 11098: contig of 2903 bp in length
* 11099
* 11198: gap of unknown length
* 11199
* 14994: contig of 3796 bp in length
* 14995
* 15094: gap of unknown length
* 15095
* 17463: contig of 2369 bp in length
* 17464
* 17563: gap of unknown length
* 17564
* 20734: contig of 3111 bp in length
* 20735
* 20834: gap of unknown length
* 23622: contig of 2788 bp in length
* 23623
* 23722: gap of unknown length
* 23723
* 27296: contig of 3574 bp in length
* 27297
* 27396: gap of unknown length

```

```

* 27397 30420: contig of 3024 bp in length
* 30421 30520: gap of unknown length
* 30521 32694: contig of 2174 bp in length
* 32695 32794: gap of unknown length
* 32795 34990: contig of 2196 bp in length
* 34991 35090: gap of unknown length
* 35091 36752: contig of 1662 bp in length
* 36753 36852: gap of unknown length
* 36853 38346: contig of 1494 bp in length
* 38347 38447: gap of unknown length
* 38448 39963: contig of 1517 bp in length
* 39964 40063: gap of unknown length
* 40064 42411: contig of 2348 bp in length
* 42412 42511: gap of unknown length
* 42512 43761: contig of 1250 bp in length
* 43762 43861: gap of unknown length
* 43862 45694: contig of 1833 bp in length
* 45695 45794: gap of unknown length
* 45795 47360: contig of 1566 bp in length
* 47361 47460: gap of unknown length
* 47461 49513: contig of 2053 bp in length
* 49514 50939: gap of unknown length
* 50940 51039: gap of unknown length
* 51040 53241: contig of 2202 bp in length
* 53242 53341: gap of unknown length
* 53342 54820: contig of 1479 bp in length
* 54821 54921: gap of unknown length
* 54922 56544: contig of 1624 bp in length
* 56545 56644: gap of unknown length
* 56645 58011: contig of 1367 bp in length
* 58012 58111: gap of unknown length
* 58112 59934: contig of 1822 bp in length
* 59935 60033: gap of unknown length
* 60034 62107: contig of 2074 bp in length
* 62108 62207: gap of unknown length
* 62209 63309: contig of 1102 bp in length
* 63310 64536: contig of 1127 bp in length
* 64537 64636: gap of unknown length
* 64637 65952: contig of 1316 bp in length
* 65953 66052: gap of unknown length
* 66053 67993: contig of 1941 bp in length
* 67994 68093: gap of unknown length
* 68094 69302: contig of 1209 bp in length
* 69303 71020: gap of unknown length
* 71021 71120: contig of 1618 bp in length
* 71121 72909: contig of 1789 bp in length
* 72910 73009: gap of unknown length
* 73010 74203: contig of 1200 bp in length
* 74210 74309: gap of unknown length
* 74310 75511: contig of 1202 bp in length
* 75512 76611: gap of unknown length
* 76612 76866: contig of 1255 bp in length
* 76867 76966: gap of unknown length
* 76967 78063: contig of 1097 bp in length
* 78064 78163: gap of unknown length
* 78164 79266: contig of 1103 bp in length
* 79267 79366: gap of unknown length
* 79367 80909: contig of 1543 bp in length
* 80910 82368: gap of unknown length
* 82369 82468: gap of unknown length
* 82469 84038: contig of 1570 bp in length
* 84039 84138: gap of unknown length
* 84139 85484: contig of 1346 bp in length
* 85485 85584: gap of unknown length
* 85585 86606: contig of 1022 bp in length.

FEATURES
Source
1. 86606 Location/Qualifiers
/organism="Rattus norvegicus"
/db_xref="taxon:10116"

```

```

BASE COUNT 23615 a 17813 c 17427 g 23494 t 4257 others
ORIGIN

```

```

Query Match 2.8%; Score 21; DB 2; Length 86606;
Best Local Similarity 100.0%; Pred.No. 1.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 674 CTTACACCTTCTGTGAGATT 694
DB 83713 CTTACACCTTCTGTGAGATT 83693

```

```

RESULT 7
LMFLCHR34_07/c
WPCOMMENT

```

```

Sequence split into 18 fragments LOCUS LMFLCHR34 Accession AL499623

```

Fragment Name	Begin	End
LMFLCHR34_00	1	11000
LMFLCHR34_01	100001	21000
LMFLCHR34_02	200001	31000
LMFLCHR34_03	300001	41000
LMFLCHR34_04	400001	51000
LMFLCHR34_05	500001	61000
LMFLCHR34_06	600001	71000
LMFLCHR34_07	700001	81000
LMFLCHR34_08	800001	91000
LMFLCHR34_09	900001	101000
LMFLCHR34_10	1000001	111000
LMFLCHR34_11	1100001	121000
LMFLCHR34_12	1200001	131000
LMFLCHR34_13	1300001	141000
LMFLCHR34_14	1400001	151000
LMFLCHR34_15	1500001	161000
LMFLCHR34_16	1600001	171000
LMFLCHR34_17	1700001	170777

Continuation (8 of 18) of LMFLCHR34 from base 700001 (AL499623 Leishmania major chrom

```

Query Match 2.8%; Score 21; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred.No. 1.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 199 CCGTCGTGCTTGCCAGCA 219
DB 70942 CCGTCGTGCTTGCCAGCA 70922

```

```

RESULT 8
LOCUS
DEFINITION Homo sapiens chromosome 16 clone RP11-309014, WORKING DRAFT
ACCESSION AC009066
VERSION AC009066.5 GI:9256122
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ADDITIONAL INFORMATION:
DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished
2 (bases 1 to 152815)
DOE Joint Genome Institute.
Direct Submission
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 18, 2000 this sequence version replaced gi:7689895.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

```

 Project Information
 Center Project Name: 525371
 Center clone name: RPCI-11_309014

Summary Statistics
 Consensus quality: 140613 bases at least Q40
 Consensus quality: 148582 bases at least Q30
 Consensus quality: 150226 bases at least Q20
 Estimated insert size: 160590; agarose-*rf* estimation
 Estimated insert size: 152165; sum-of-*contigs* estimation
 Quality coverage: 6.7 in Q20 bases; agarose-*rf* estimation
 Quality coverage: 7.08 in Q20 bases; sum-of-*contigs* estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 15 *contigs*. Gaps between the *contigs*
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 1 1799: *contig* of 1799 bp in length
 * 1800 1899: *gap* of unknown length
 * 1900 4720: *contig* of 2821 bp in length
 * 4721 4820: *gap* of unknown length
 * 4821 7854: *contig* of 3034 bp in length
 * 7855 7954: *gap* of unknown length
 * 7955 11592: *contig* of 3638 bp in length
 * 11593 11692: *gap* of unknown length
 * 11693 14169: *contig* of 2477 bp in length
 * 14170 14269: *gap* of unknown length
 * 14270 28549: *contig* of 14280 bp in length
 * 28550 28649: *gap* of unknown length
 * 28650 52060: *contig* of 23411 bp in length
 * 52061 52160: *gap* of unknown length
 * 52161 107022: *contig* of 54862 bp in length
 * 107023 107122: *gap* of unknown length
 * 107123 109610: *contig* of 2488 bp in length
 * 109611 109710: *gap* of unknown length
 * 109711 133288: *contig* of 23578 bp in length
 * 133289 133388: *gap* of unknown length
 * 133389 134596: *contig* of 1208 bp in length
 * 134597 134696: *gap* of unknown length
 * 134697 136172: *contig* of 1476 bp in length
 * 136173 136272: *gap* of unknown length
 * 136273 146371: *contig* of 10099 bp in length
 * 146372 146471: *gap* of unknown length
 * 146472 151105: *contig* of 4534 bp in length
 * 151006 151105: *gap* of unknown length
 * 151106 152815: *contig* of 1710 bp in length.

FEATURES
 source
 1..152815
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-309014"
 /clone_lib="RPCI human BAC library 11"

BASE COUNT 40215 a 31837 c 33501 g 45857 t 1405 others
 ORIGIN

Query Match 2.8%; Score 21; DB 2; Length 152815;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 441 ATGCTGAGTGGCTGTTGGGG 461
 ||||||||||||||||||
 DB 100610 ATGCTGAGTGGCTGTTGGGG 100630

RESULT 9
 AC096247/c AC096247 159759 bp DNA linear HTG 20-DEC-2001
 LOCUS

DEFINITION Rattus norvegicus clone CH230-30J6. *** SEQUENCING IN PROGRESS ***
 74 unordered pieces.

AC096247
 AC096247.2 GI:17943941
 HTG: HTGS_PHASE1
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., All-ouman,F.R., Allen,C.,
 Alibabak,S.L., Amaralungu,H.C., Aye,U.R., Banks,T., Barbata,J.,
 Benton,J., Bimaga,K., Blankenbiller,K., Bonin,D., Bouck,J.,
 Bowler,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buha,C.,
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
 Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
 Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
 Denn,A.L., Ding,Y., Dinh,H.H., Douhalwaite,K.J., Draper,H.,
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
 Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flaggs,N., Ford,J.,
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
 Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S.,
 Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
 Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
 Hollins,B., Homsli,F., Howard,S., Huber,J., Huijck,S., Hume,J.,
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korah,J.,
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
 Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,D., Liu,W.,
 Louisedge,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
 Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
 Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabadi,K.,
 Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
 Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Norkenko,S.,
 Ogihara,M., Okuwon,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.,
 Qulles,M., Ren,Y., Rivers,M., Rojas,A., Rojoubkan,I., Rolfe,M.,
 Ruiz,S., Savary,G., Scherer,S., Scott,G., Sher,H., Shoshitari,N.,
 Sisson,I., Sodergren,E., Sonstake,T., Sparks,A., Stanley,H.,
 Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellor,B., Thomas,N.,
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalona,D., Vinson,R.,
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
 Wellington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S.,
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G., and Gibbs,R.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT
 Direct Submission
 2 (bases 1 to 159759)
 Unpublished
 Direct Submission
 Worley,K.C.
 Direct Submission
 Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Dec 20, 2001 this sequence version replaced gi:15627867.

 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

 Project Information
 Center project name: GFTL
 Center clone name: CH230-30J6

 Assembly program: Phrap; version 0.990329first call to
 findPhraplist
 Consensus quality: 121344 bases at least Q40
 Consensus quality: 131650 bases at least Q30
 Consensus quality: 140303 bases at least Q20

Estimated insert size: 121638; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 1.9x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 74 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

```
1 6091: contig of 6091 bp in length
* 6092 6191: gap of unknown length
* 6192 13708: contig of 7517 bp in length
* 13709 13808: gap of unknown length
* 13809 18731: contig of 4923 bp in length
* 18732 18831: gap of unknown length
* 18832 23050: contig of 4219 bp in length
* 23051 23150: gap of unknown length
* 23151 26797: contig of 3647 bp in length
* 26798 26897: gap of unknown length
* 26898 30055: contig of 3158 bp in length
* 30056 30155: gap of unknown length
* 30156 34216: contig of 4061 bp in length
* 34217 34316: gap of unknown length
* 34317 38691: contig of 4375 bp in length
* 38692 41706: contig of 2915 bp in length
* 41707 41806: gap of unknown length
* 41807 45279: contig of 3473 bp in length
* 45280 45379: gap of unknown length
* 45380 48515: contig of 3136 bp in length
* 48516 51208: gap of unknown length
* 51209 51308: contig of 2593 bp in length
* 51309 54184: gap of unknown length
* 54185 54284: gap of unknown length
* 54285 56323: contig of 2039 bp in length
* 56324 59464: gap of unknown length
* 59465 59564: contig of 3041 bp in length
* 59565 63185: gap of unknown length
* 63186 63285: gap of unknown length
* 63286 65883: contig of 2498 bp in length
* 65884 65806: gap of unknown length
* 65807 68606: gap of unknown length
* 68607 70680: gap of unknown length
* 70681 70780: contig of 2074 bp in length
* 70781 71979: gap of unknown length
* 71980 72079: gap of unknown length
* 72080 74302: contig of 2223 bp in length
* 74303 74402: gap of unknown length
* 74403 75732: contig of 1330 bp in length
* 75733 75832: gap of unknown length
* 75833 77000: contig of 1168 bp in length
* 77001 77100: gap of unknown length
* 77101 78964: contig of 1864 bp in length
* 78965 79064: gap of unknown length
* 79065 81524: contig of 2460 bp in length
* 81525 81624: gap of unknown length
* 81625 83011: contig of 1387 bp in length
* 83012 85356: gap of unknown length
* 85357 85456: contig of 2245 bp in length
* 85457 87032: gap of unknown length
* 87033 87132: contig of 1576 bp in length
* 87133 88850: gap of unknown length
* 88851 88950: contig of 1718 bp in length
* 88951 90855: contig of 1905 bp in length
```

```
90856 90955: gap of unknown length
* 90956 92017: contig of 1062 bp in length
* 92018 92117: gap of unknown length
* 92118 93487: contig of 1370 bp in length
* 93488 93587: gap of unknown length
* 93588 95224: contig of 1637 bp in length
* 95225 95324: gap of unknown length
* 95325 97110: contig of 1786 bp in length
* 97111 97211: gap of unknown length
* 97212 99100: contig of 1890 bp in length
* 99101 99201: gap of unknown length
* 99202 100888: contig of 1688 bp in length
* 100889 100988: gap of unknown length
* 100989 102438: contig of 1450 bp in length
* 102439 102538: gap of unknown length
* 102539 104343: contig of 1805 bp in length
* 104344 104443: gap of unknown length
* 104444 106443: contig of 2000 bp in length
* 106444 108915: gap of unknown length
* 108916 109015: contig of 2372 bp in length
* 109016 110825: gap of unknown length
* 110826 110925: contig of 1810 bp in length
* 110926 112773: contig of 1848 bp in length
* 112774 112873: gap of unknown length
* 112874 114751: contig of 1878 bp in length
* 114752 114851: gap of unknown length
* 114852 116307: contig of 1456 bp in length
* 116308 116407: gap of unknown length
* 116408 117866: contig of 1459 bp in length
* 117867 117966: gap of unknown length
* 117967 119525: contig of 1559 bp in length
* 119526 119625: gap of unknown length
* 119626 121097: contig of 1472 bp in length
* 121098 121197: gap of unknown length
* 121198 122331: contig of 1134 bp in length
* 122332 122431: gap of unknown length
* 122432 123782: contig of 1351 bp in length
* 123783 123882: gap of unknown length
* 123883 126212: contig of 2330 bp in length
* 126213 126312: gap of unknown length
* 126313 127601: contig of 1289 bp in length
* 127602 127701: gap of unknown length
* 127702 128877: contig of 1176 bp in length
* 128878 128977: gap of unknown length
* 128978 130350: contig of 1373 bp in length
* 130351 130450: gap of unknown length
* 130451 131571: contig of 1121 bp in length
* 131572 131671: gap of unknown length
* 131672 132866: contig of 1195 bp in length
* 132867 132966: gap of unknown length
* 132967 134526: contig of 1560 bp in length
* 134527 134626: gap of unknown length
```

Query Match 2.8% Score 21; DB 2; Length 159759;
Best Local Similarity 100.0%; Pred. No. 1.8;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 CTGCTTGAGCAGTGTGCCC 271
Db 157980 CTGCTTGAGCAGTGTGCCC 157960

RESULT 10
AC026963/c 178954 bp. DNA 1linear HTG 03-JUL-2000
LOCUS Homo sapiens chromosome 18 clone RP11-294A3 map 18, WORKING DRAFT
DEFINITION AC026963
SEQUENCE AC026963
ACCESSION AC026963.2 GI:8901252
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 178954)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 18, clone RP11-294A3
Unpublished
2 (bases 1 to 178954)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barra, N., Bastien, V., Beda, F., Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G., Campolano, A., Castle, A., Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Deatrelano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Harford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Larocque, K., Lamazares, R., Landers, T., Lehocsky, J., Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Mhova, T., Miranda, C., Mienga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisanzi, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Rhmann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, V., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (25-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 3, 2000 this sequence version replaced 91.732842.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: 294_A3

Center clone name: L8505

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 164858 bases at least Q40

Consensus quality: 171910 bases at least Q30

Consensus quality: 174306 bases at least Q20

Insert size: 180000; agarose-fp

Insert size: 175854; sum-of-coverage

Quality coverage: 4.3 in Q20 bases; agarose-fp

Quality coverage: 4.4 in Q20 bases; sum-of-coverage

* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1260: contig of 1260 bp in length
* 1261 1360: gap of 100 bp
* 1361 2401: contig of 1041 bp in length
* 2402 2501: gap of 100 bp
* 2502 4093: contig of 1592 bp in length
* 4094 4193: gap of 100 bp
* 4194 5657: contig of 1464 bp in length
* 5658 5757: gap of 100 bp
* 5758 7508: contig of 1751 bp in length
* 7509 7608: gap of 100 bp

FEATURES
SOURCE

7609 9930: contig of 2322 bp in length
9931 10030: gap of 100 bp
10031 11954: contig of 1924 bp in length
11955 12054: gap of 100 bp
12055 15202: contig of 3148 bp in length
15203 15302: gap of 100 bp
15303 17660: contig of 2358 bp in length
17661 17760: gap of 100 bp
17761 20163: contig of 2403 bp in length
20164 20263: gap of 100 bp
20264 23557: contig of 3294 bp in length
23558 23657: gap of 100 bp
23658 25641: contig of 1984 bp in length
25642 25741: gap of 100 bp
25742 28811: contig of 3070 bp in length
28812 28911: gap of 100 bp
28912 33117: contig of 4206 bp in length
33118 33217: gap of 100 bp
33218 36910: contig of 3693 bp in length
36911 37010: gap of 100 bp
37011 40541: contig of 3531 bp in length
40542 40641: gap of 100 bp
40642 45984: contig of 5343 bp in length
45985 46084: gap of 100 bp
46085 51778: contig of 5694 bp in length
51779 51878: gap of 100 bp
51879 56627: contig of 4749 bp in length
56628 56727: gap of 100 bp
56728 64204: contig of 7477 bp in length
64205 64304: gap of 100 bp
64305 71466: contig of 7162 bp in length
71467 71566: gap of 100 bp
71567 78191: contig of 6625 bp in length
78192 78291: gap of 100 bp
78292 86740: contig of 8449 bp in length
86741 86840: gap of 100 bp
86841 94265: contig of 7425 bp in length
94266 94365: gap of 100 bp
94366 101783: contig of 7418 bp in length
101784 101883: gap of 100 bp
101884 108250: contig of 6367 bp in length
108251 108350: gap of 100 bp
108351 119190: contig of 10840 bp in length
119191 119290: gap of 100 bp
119291 129743: contig of 10453 bp in length
129744 129843: gap of 100 bp
129844 138898: contig of 9055 bp in length
138899 138998: gap of 100 bp
138999 151653: contig of 12655 bp in length
151654 151753: gap of 100 bp
151754 164959: contig of 13206 bp in length
164960 165059: gap of 100 bp
165060 178954: contig of 13895 bp in length.
Location/Qualifiers
1. 178954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-294A3"
/clone_lib="RP11 Human Male BAC"
1. 1260
/note="assembly-fragment"
1361. 2401
/note="assembly-fragment"
2502. 4093
/note="assembly-fragment"
4194. 5657
/note="assembly-fragment"
5758. 7508
/note="assembly-fragment"
7609. 9930
/note="assembly-fragment"

```

misc_feature 10031..11954
                /note="assembly-fragment"
misc_feature 12055..15202
                /note="assembly-fragment"
misc_feature 15303..17660
                /note="assembly-fragment"
misc_feature 17661..20163
                /note="assembly-fragment"
misc_feature 20264..23557
                /note="assembly-fragment"
                /note="assembly-fragment"
                clone_end:"7"
                vector_side:left"
misc_feature 23558..25641
                /note="assembly-fragment"
misc_feature 25742..28811
                /note="assembly-fragment"
misc_feature 28912..33117
                /note="assembly-fragment"
misc_feature 33218..36910
                /note="assembly-fragment"
misc_feature 37011..40541
                /note="assembly-fragment"
misc_feature 40642..45984
                /note="assembly-fragment"
misc_feature 46085..51778
                /note="assembly-fragment"
misc_feature 51879..55627
                /note="assembly-fragment"
misc_feature 56728..64204
                /note="assembly-fragment"
misc_feature 64305..71466
                /note="assembly-fragment"
misc_feature 71567..78191
                /note="assembly-fragment"
                clone_end:Sp6
                vector_side:right"
misc_feature 78292..86740
                /note="assembly-fragment"
misc_feature 86841..94265
                /note="assembly-fragment"
misc_feature 94366..101783
                /note="assembly-fragment"
misc_feature 101884..108250
                /note="assembly-fragment"
misc_feature 108351..119190
                /note="assembly-fragment"

Query Match      2.8%: Score 21; DB 2; Length 178954;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 441 ATGCTGAGTGCTGTTGGG 461
      |||||
Db 61743 ATGCTGAGTGCTGTTGGG 61723

RESULT 11
AC103584      179982 bp   DNA      linear   HTG 29-JAN-2002
LOCUS      Bos taurus clone rp42-51p13, WORKING DRAFT SEQUENCE, 11 unordered
DEFINITION      pieces
AC103584      GI:18390279
VERSION      HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS      COV.
SOURCE      Bos taurus
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
                Bovidae; Bovinae; Bos.
REFERENCE      1 (bases 1 to 179982)
AUTHORS      Lau,C.C. and Roe,B.A.
TITLE      Bos taurus BAC Clone rp42-51p13
JOURNAL      Unpublished

```

```

REFERENCE      2 (bases 1 to 179982)
AUTHORS      Lau,C.C. and Roe,B.A.
TITLE      Direct Submission
JOURNAL      Submitted (29-NOV-2001) Department of Chemistry And Biochemistry,
                The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
                OK 73019, USA
COMMENT      On Jan 29, 2002 this sequence version replaced g1.18139437.
                ----- Genome Center
                Center: Department Of Chemistry And Biochemistry
                The University Of Oklahoma
                Center code:UOKNOR
                -----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2012: contig of 2012 bp in length
* 2013 2112: gap of unknown length
* 2113 6595: contig of 4483 bp in length
* 6596 6695: gap of unknown length
* 6696 10370: contig of 3675 bp in length
* 10371 10470: gap of unknown length
* 10471 15087: contig of 4617 bp in length
* 15088 15187: gap of unknown length
* 15188 20215: contig of 5028 bp in length
* 20216 20315: gap of unknown length
* 20316 31332: contig of 10917 bp in length
* 31333 31332: gap of unknown length
* 31333 45692: contig of 14360 bp in length
* 45693 45792: gap of unknown length
* 45793 62170: contig of 16378 bp in length
* 62171 62270: gap of unknown length
* 62271 79892: contig of 17622 bp in length
* 79893 79992: gap of unknown length
* 79993 119784: contig of 39792 bp in length
* 119785 119884: gap of unknown length
* 119885 179982: contig of 60098 bp in length.
                Location/Qualifiers
                1..179982
                /organism="Bos taurus"
                /db_xref="taxon:9913"
                /clone="rp42-51p13"

BASE COUNT      46071 a 43236 c 41984 g 47663 t 1028 others
ORIGIN

Query Match      2.8%: Score 21; DB 2; Length 179982;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 TCTCAGGCTCCTCTGCTCAC 90
      |||||
Db 65193 TCTCAGGCTCCTCTGCTCAC 65173

RESULT 12
AC007225      218892 bp   DNA      linear   PRI 19-JAN-2000
LOCUS      Homo sapiens chromosome 16 clone RCT-11_48057, complete sequence.
DEFINITION      AC007225
AC007225      GI:6715703
VERSION      HTG.
KEYWORDS      HTG.
SOURCE      human.
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1 (bases 1 to 218892)
AUTHORS      Bruce,D., Mundt,M., Doggett,N., Munk,C., Saunders,E., Robinson,D.,
                Jones,M., Bucklingham,J., Chasteen,L., Thompson,S., Goodwin,L.,
                Bryant,J., Tesmer,J., Meincke,L., Longmire,J., White,S., Tatum,O.,

```

```
Campbell,C., Fawcett,J., Maltbie,M., Bussod,M., Sutherland,R.,
Han,C. and Deaven,L.
TITLE
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 218892)
AUTHORS
Ricke,D.O.
TITLE
Large Scale Sequence Analysis and Annotation with the Sequence
Comparison Analysis (SCAN) System
JOURNAL
Unpublished
REFERENCE
3 (bases 1 to 218892)
AUTHORS
Bruce,D., Mundt,M., Doggett,N., Munk,C., Saunders,E., Robinson,D.,
Jones,M., Buckingham,J., Chasteen,L., Thompson,S., Goodwin,L.,
Bryant,J., Tesmer,J., Melnick,L., Longmire,J., White,S., Tatum,O.,
Campbell,C., Fawcett,J., Maltbie,M., Bussod,M., Sutherland,R.,
McMurry,K., Han,C. and Deaven,L.
TITLE
Direct Submission
JOURNAL
Submitted (06-APR-1999) Center for Human Genome Studies, DOE Joint
Genome Institute, Los Alamos National Laboratory, MS M888, Los
Alamos, NM 87545, USA
REFERENCE
4 (bases 1 to 218892)
AUTHORS
Bruce,D., Mundt,M., Doggett,N., Munk,C., Saunders,E., Robinson,D.,
Jones,M., Buckingham,J., Chasteen,L., Thompson,S., Goodwin,L.,
Bryant,J., Tesmer,J., Melnick,L., Longmire,J., White,S., Tatum,O.,
Campbell,C., Fawcett,J., Maltbie,M., Bussod,M., Sutherland,R.,
McMurry,K., Han,C. and Deaven,L.
TITLE
Direct Submission
JOURNAL
Submitted (19-JAN-2000) Center for Human Genome Studies, DOE Joint
Genome Institute, Los Alamos National Laboratory, MS M888, Los
Alamos, NM 87545, USA
COMMENT
On Jan 19, 2000 this sequence version replaced gi:567163.
FEATURES
Location/Qualifiers
source
1..218892
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RPC1-11_48067"
/complement(589..853)
/note="GRAIL 2 excellent exon, frame 1"
/complement(1526..1623)
/note="GRAIL 2 excellent exon, frame 0"
/complement(3450..3774)
/rpt_family="Alu"
repeat_region
5258..5353
/rpt_family="Alu"
repeat_region
5615..5747
/rpt_family="Alu"
repeat_region
complement(6186..6348)
/rpt_family="MERS"
repeat_region
6576..6652
/rpt_family="Alu"
repeat_region
6666..6977
/rpt_family="L1"
repeat_region
7063..7149
/rpt_family="Alu"
repeat_region
7139..7283
/rpt_family="L1"
repeat_region
/rpt_family="Alu"
complement(7364..7656)
/rpt_family="Alu"
repeat_region
complement(8004..8501)
/rpt_family="Alu"
repeat_region
/rpt_family="Alu"
complement(9329..9630)
/rpt_family="Alu"
repeat_region
/rpt_family="Alu"
complement(9671..9990)
/rpt_family="Alu"
repeat_region
complement(9919..10260)
/rpt_family="Alu"
repeat_region
complement(10375..10556)
/rpt_family="L1"
repeat_region
complement(10602..10867)
/rpt_family="Alu"
repeat_region
complement(10864..11169)
/rpt_family="Alu"
repeat_region
complement(11158..11365)
/rpt_family="Alu"
repeat_region
complement(11449..11512)
/rpt_family="MIR"
repeat_region
11540..11810
/rpt_family="Alu"
repeat_region
complement(12445..12839)
/rpt_family="Alu"
repeat_region
13055..13854
/rpt_family="THE1"
repeat_region
13106..13787
/rpt_family="MS7AR"
repeat_region
complement(13885..14267)
/rpt_family="Alu"
repeat_region
complement(14409..14787)
/rpt_family="Alu"
repeat_region
complement(15068..15155)
/rpt_family="Alu"
repeat_region
15164..15339
/rpt_family="THE1"
repeat_region
complement(15425..15798)
/rpt_family="Alu"
repeat_region
complement(15854..16087)
/rpt_family="Alu"
repeat_region
16103..16179
/rpt_family="THE1"
repeat_region
16182..17770
/rpt_family="MS7AR"
repeat_region
17771..17910
/rpt_family="THE1"
repeat_region
complement(17968..18256)
/rpt_family="Alu"
repeat_region
18252..18477
/rpt_family="THE1"
repeat_region
complement(18506..18805)
/rpt_family="Alu"
repeat_region
complement(18992..19158)
/rpt_family="Alu"
repeat_region
complement(19616..19897)
/rpt_family="Alu"
repeat_region
19952..20249
/rpt_family="Alu"
repeat_region
complement(20373..20677)
/rpt_family="Alu"
repeat_region
complement(20940..21227)
/rpt_family="Alu"
repeat_region
complement(21743..21957)
/rpt_family="Alu"
repeat_region
22226..22479
/rpt_family="Alu"
repeat_region
22834..23194
/rpt_family="Alu"
repeat_region
complement(23930..24226)
/rpt_family="Alu"
repeat_region
complement(24227..24536)
/rpt_family="Alu"
repeat_region
25671..26079
/rpt_family="Alu"
repeat_region
complement(26081..26376)
/rpt_family="Alu"
repeat_region
26552..26596
/rpt_family="MIR"
repeat_region
27005..27290
/rpt_family="MER"
repeat_region
complement(28321..28625)
/rpt_family="Alu"
repeat_region
28663..28958
/rpt_family="Alu"
repeat_region
complement(29011..29310)
/rpt_family="Alu"
repeat_region
30537..30845
/rpt_family="Alu"
repeat_region
31696..31963
/rpt_family="Alu"
repeat_region
```


Db 47 GAGATGCCGTGCTGATGG 28

RESULT 15					
ATACOC/C	6605 bp	DNA	linear	PLN 05-NOV-1997	
LOCUS	ATACOC				
DEFINITION	A.thaliana gene encoding acetyl-CoA carboxylase, biotin subunit.				
ACCESSION	y09061				
VERSION	y09061.1	GI:1865670			
KEYWORDS	acetyl-CoA carboxylase; biotin carboxylase;				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.				
REFERENCE	1 (bases 1 to 6605)				
AUTHORS	Bao,X., Shorrosh,B.S. and Ohlrogge,J.B.				
TITLE	Isolation and characterization of an Arabidopsis biotin carboxylase				
	gene and its promoter				
JOURNAL	Plant Mol. Biol. 35 (5), 539-550 (1997)				
MEDLINE	98009989				
REFERENCE	2 (bases 1 to 6605)				
AUTHORS	Ohlrogge,J.B.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-OCT-1996) J.B. Ohlrogge, Michigan State University,				
	Dept. Of Botany and Plant Pathology, East Lansing, MI 48824-1312,				
	USA				

TATA_signal	2021.. 2027
exon	10101,2078 . 2319,2417 . 2760,2838 . 2979,3119 . 3217,3309 . 3365,3569 . 3640,4007 . 4096,4232 . 4291,4597 . 4864 . 4956,5035 . 5116,5205 . 5311,5527 . 5602,5757 . 5803,5892 . 5954,6027 . 6083)
intron	2078 . 2319
exon	/number=1
intron	10101,2272 . 2319,2417 . 2760,2838 . 2979,3119 . 3217,3309 . 3365,3569 . 3640,4007 . 4096,4232 . 4291,4597 . 4864 . 4956,5035 . 5116,5205 . 5311,5527 . 5602,5757 . 5803,5892 . 5954,6027 . 6083)
exon	/EC_number="6.4.1.2"
intron	/codon_start=1
exon	/product="acetyl-CoA carboxylase, biotin carboxylase subunit"
intron	/protein_id="CAA70282.1"
exon	/db_xref="GI:1865671"
intron	/db_xref="SPTREMBL:P33650"
exon	/translation="WDASMITNSKTSITSPSLAGKSGGGVIRSSLCNIMPSKYNFPROPIRQTSOKKIKRATISGELGVTCGSGKILVANGRTAVYITAEIMQIPCYAALVSTFDKALHVLADAEVCIAGEASNOSYVLPNTVAISGCTMLHGVEFLEENALFVCMGDKHINFLTGPMPDPIRYMGDKATNREMKNAQIPYVPGSGILQSTDEAVR
intron	VANEGPFWMTATAGGCGGMRLAKEGEVVKLLQAKSEAAAFGNDCCYLEKTVQ
exon	NPRHLEIVLADKFGNVVHFEGRDCSTQRRNOKLLEPSPALAEALKAGDAVAA
intron	ASIGIGIVGVEFLDKRSGFVEMEMNTRIQVEHPEVEMISVDLIEQIRVMAGEEL
exon	RKYQDILVIRGHSIECRINADPKRFGPGGRTSYLPSGGPGRMDSHYSDYVVP
intron	PSYDLIGKLTIVMAPTREKATREMKRRLANPTITIGVPTTTINRKLIDVDEPKNQVD
exon	TAFYIKHEEELAEVLGVTEAFGAKAEHRDRIA"
intron	2320 . 2416
exon	/number=1
intron	2417 . 2760
exon	/number=2
intron	2761 . 2837
exon	/number=2
intron	2838 . 2979
exon	/number=3
intron	2980 . 3118
exon	/number=3
intron	3119 . 3217
exon	/number=4

BASE ORIGIN	COUNT	2106	a	1068	c	1197	g	2201	t	33	others
polya_signal	6492	6497									
exon	6027	6083									
intron	4097	4231									
exon	4232	4291									
intron	4292	4596									
exon	4597	4779									
intron	4780	4863									
exon	4864	4956									
intron	4957	5034									
exon	5035	5116									
intron	5117	5204									
exon	5205	5311									
intron	5312	5526									
exon	5527	5602									
intron	5603	5756									
exon	5757	5803									
intron	5804	5891									
exon	5892	5954									
intron	5955	6026									
exon	6027	6083									
intron	6084	6170									
exon	6171	6257									
intron	6258	6344									
exon	6345	6431									
intron	6432	6518									
exon	6519	6605									
intron	6606	6692									
exon	6693	6779									
intron	6780	6866									
exon	6867	6953									
intron	6954	7040									
exon	7041	7127									
intron	7128	7214									
exon	7215	7301									
intron	7302	7388									
exon	7389	7475									
intron	7476	7562									
exon	7563	7649									
intron	7650	7736									
exon	7737	7823									
intron	7824	7910									
exon	7911	8000									
intron	8001	8090									
exon	8091	8180									
intron	8181	8270									
exon	8271	8360									
intron	8361	8450									
exon	8451	8540									
intron	8541	8630									
exon	8631	8720									

Search completed: November 3, 2002, 02:56:39
Job time : 3212 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 23:03:22 ; Search time 1822 Seconds

(Without alignments)
8763.418 Million cell updates/sec

Title: US-09-092-297-5

Sequence: 1 CTCACCTGCAACACACGAGG.....ATAATTATGCTACTTTATA 763

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_hg_hum.*
31: em_hg_inv.*
32: em_hg_other.*
33: em_hgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query
No. Score Match Length DB ID Description

1	761	99.7	771	6	AX164131	AX164131 Sequence
2	760.6	99.7	762	6	AX014140	AX014140 Sequence
3	418	54.8	107573	9	AC006262	AC006262 Homo sapi
4	140.6	18.4	146180	9	AC007193	AC007193 Homo sapi
5	90.6	11.9	662	6	AX055440	AX055440 Homo sapi
6	90.6	11.9	662	6	AX055696	AX055696 Sequence
7	88.8	11.6	378	6	AX118905	AX118905 Sequence
8	70.2	9.2	166500	9	AC007785	AC007785 Homo sapi
9	64.8	8.5	166500	9	AC007785	AC007785 Homo sapi
10	47.2	6.2	125020	9	AF429315	AF429315 Homo sapi
11	40.8	5.3	110000	2	AC013622_0	AC013622 Mus muscu
12	40.8	5.3	110000	2	AC013622_1	AC013622 Mus muscu
13	40.8	5.3	110000	2	AC021632_0	AC021632 Mus muscu
14	40.8	5.3	207588	2	AC087152	AC087152 Homo sapi
15	40.4	5.3	172090	2	AC018477	AC018477 Homo sapi
16	40.4	5.3	200430	9	AC011500	AC011500 Homo sapi
17	39.8	5.2	10732	6	E32986	E32986 Gene encod
18	39.4	5.2	125020	9	AF429315	AF429315 Homo sapi
19	39.4	5.2	138134	2	AC009718	AC009718 Mus muscu
20	39.4	5.2	156958	2	AC015869	AC015869 Homo sapi
21	38.8	5.1	239147	2	AC094982	AC094982 Rattus no
22	38.4	5.0	5067	2	AB077269	AB077269 Desulfovi
23	38.2	5.0	207187	2	AC012555	AC012555 Homo sapi
24	38.2	5.0	237252	2	AC099543	AC099543 Homo sapi
25	37.6	4.9	95963	2	AC106422	AC106422 Rattus no
26	37.2	4.9	158274	2	AC107622	AC107622 Homo sapi
27	37	4.8	1398	10	AF106278	AF106278 Mus muscu
28	37	4.8	1511	10	MMU316613	MMU316613 Mus muscu
29	37	4.8	1534	10	AF043943	AF043943 Mus muscu
30	37	4.8	2636	10	BC015074	BC015074 Mus muscu
31	37	4.8	102739	9	HS671014	HS671014 Mus muscu
32	37	4.8	137346	2	AP003622	AP003622 Human DNA
33	37	4.8	166706	2	HSB42686	HSB42686 Homo sapi
34	36.8	4.8	175137	2	AC018670	AC018670 Homo sapi
35	36.6	4.8	173660	9	AC084693	AC084693 Homo sapi
36	36.6	4.8	176708	9	AC039056	AC039056 Homo sapi
37	36.4	4.8	72847	2	AC023563	AC023563 Homo sapi
38	36.4	4.8	110000	2	AC092160_1	AC092160 Homo sapi
39	36.4	4.8	139416	2	AC068963	AC068963 Homo sapi
40	36.4	4.8	150891	2	AC069098	AC069098 Homo sapi
41	36.4	4.8	167592	9	AC018636	AC018636 Homo sapi
42	36.4	4.8	185069	2	AC084692	AC084692 Homo sapi
43	36.4	4.8	188711	9	AC093326	AC093326 Homo sapi
44	36.4	4.8	211564	2	AC092159	AC092159 Homo sapi
45	36.4	4.8	246838	2	AC008026	AC008026 Homo sapi

ALIGNMENTS

RESULT 1
AX164131
LOCUS AX164131
DEFINITION Sequence 3 from Patent WO0140465.
ACCESSION AX164131
VERSION AX164131.1 GI:14545080
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS Rong,S., Goddard,A., Godowski,P.J., Grimaldi,C.J., Gurney,A.L.,
Hillan,K.J., Tumas,D., Watanabe,C.K., Wood,W.I. and Zhang,Z.
TITLE Compositions and methods for the treatment of immune related
diseases
JOURNAL Patent: WO 0140465-A 3 07-JUN-2001;
Genentech, Inc. (US)
FEATURES
Source Location/Qualifiers
1..771
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 169 a 231 c 195 g 176 t

ORIGIN

Query Match 99.7%; Score 761; DB 6; Length 771;
 Best Local Similarity 99.7%; Pred. No. 1.2e-197;
 Matches 761; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 CTCACATGCAACCAACCAAGAGCCATGCTCCCGAGGCTGCATGATGCTGTCTTTGCCA 60
 |||||||
 1 CTCACATGCAACCAACCAAGAGCCATGCTCCCGAGGCTGCATGATGCTGTCTTTGCCA 60
 |||||||
 61 TTTTCTCATCTCCAGGCTCTCTGCTCACAAGGAGCCCAAGTGGCCCATGACTCTT 120
 |||||||
 61 TTTTCTCATCTCCAGGCTCTCTGCTCACAAGGAGCCCAAGTGGCCCATGACTCTT 120
 |||||||
 121 ACCGATGCTGTGGCAGGACACAGAGATGTGGGAGCAAGTTTACAGCCCTGCGAGC 180
 |||||||
 121 ACCGATGCTGTGGCAGGACACAGAGATGTGGGAGCAAGTTTACAGCCCTGCGAGC 180
 |||||||
 181 ACTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 |||||||
 181 ACTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 |||||||
 241 CCTTCAGAGTCTGCTTTGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 |||||||
 241 CCTTCAGAGTCTGCTTTGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 |||||||
 301 AGAAGTGGAGTCTGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 |||||||
 301 AGAAGTGGAGTCTGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 |||||||
 361 ACATCAGGAGGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 |||||||
 361 ACATCAGGAGGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 |||||||
 421 GGTGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 |||||||
 421 GGTGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 |||||||
 481 AACGCCCCCACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 |||||||
 481 AACGCCCCCACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 |||||||
 541 CCACAGTCTCCCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 |||||||
 541 CCACAGTCTCCCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 |||||||
 601 GACTCTGACCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 |||||||
 601 GACTCTGACCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 |||||||
 661 TGGCTGGGGAACCTTCAACCTTCTGATGATGATGATGATGATGATGATGATGAT 720
 |||||||
 661 TGGCTGGGGAACCTTCAACCTTCTGATGATGATGATGATGATGATGATGATGAT 720
 |||||||
 721 CAGGAGCAAGCAGACAGATCATATTAATTTATGACTTTATA 763
 |||||||
 721 CAGGAGCAAGCAGACAGATCATATTAATTTATGACTTTATA 763
 |||||||

RESULT 2
 AX014140 762 bp DNA linear PAT 07-SEP-2000
 LOCUS AX014140
 DEFINITION Sequence 3 from Patent WO954447.
 ACCESSION AX014140
 VERSION AX014140.1 GI:10040587
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 1 (bases 1 to 762)
 AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
 Pilarsky,C.

TITLE Human nucleic acid sequences of bladder tumour tissue

JOURNAL Patent: WO 954447-A 3 28-OCT-1999;
 SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
 BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GBS FUER GENOMFORSCHUN
 (DE); PILARSKY CHRISTIAN (DE)
 (DE); Location/Qualifiers

FEATURES

source 1.762
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 162 a 232 c 194 g 174 t

ORIGIN

Query Match 99.7%; Score 760.6; DB 6; Length 762;
 Best Local Similarity 99.9%; Pred. No. 1.6e-197;
 Matches 760; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 CTCACATGCAACCAACCAAGAGCCATGCTCCCGAGGCTGCATGATGCTGTCTTTGCCA 60
 |||||||
 1 CTCACATGCAACCAACCAAGAGCCATGCTCCCGAGGCTGCATGATGCTGTCTTTGCCA 60
 |||||||
 61 TTTTCTCATCTCCAGGCTCTCTGCTCACAAGGAGCCCAAGTGGCCCATGACTCTT 120
 |||||||
 61 TTTTCTCATCTCCAGGCTCTCTGCTCACAAGGAGCCCAAGTGGCCCATGACTCTT 120
 |||||||
 121 ACCGATGCTGTGGCAGGACACAGAGATGTGGGAGCAAGTTTACAGCCCTGCGAGC 180
 |||||||
 121 ACCGATGCTGTGGCAGGACACAGAGATGTGGGAGCAAGTTTACAGCCCTGCGAGC 180
 |||||||
 181 ACTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 |||||||
 181 ACTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 |||||||
 241 CCTTCAGAGTCTGCTTTGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 |||||||
 241 CCTTCAGAGTCTGCTTTGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 |||||||
 301 AGAAGTGGAGTCTGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 |||||||
 301 AGAAGTGGAGTCTGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 |||||||
 361 ACATCAGGAGGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 |||||||
 361 ACATCAGGAGGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 |||||||
 421 GGTGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 |||||||
 421 GGTGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 |||||||
 481 AACGCCCCCACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 |||||||
 481 AACGCCCCCACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 |||||||
 541 CCACAGTCTCCCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 |||||||
 541 CCACAGTCTCCCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 |||||||
 601 GACTCTGACCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 |||||||
 601 GACTCTGACCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 |||||||
 661 TGGCTGGGGAACCTTCAACCTTCTGATGATGATGATGATGATGATGATGATGAT 720
 |||||||
 661 TGGCTGGGGAACCTTCAACCTTCTGATGATGATGATGATGATGATGATGATGAT 720
 |||||||
 721 CAGGAGCAAGCAGACAGATCATATTAATTTATGACTTTATA 761
 |||||||
 721 CAGGAGCAAGCAGACAGATCATATTAATTTATGACTTTATA 761
 |||||||

RESULT 3
 AC006262 107573 bp DNA linear PRI 21-SEP-2001
 LOCUS AC006262
 DEFINITION Homo sapiens chromosome 19, BAC BC264576 (CTC-297N14), complete

[illegible]

QY 584 1GCTGATGGGGGCCAGGAGACTCTGAAACCTCTCTGATGATGCCCTATGGCCACAATCAAC 643
 Db 82388 TGCTGATGGGGGCCAGGAGACTCTGAAACCTCTCTGATGATGCCCTATGGCCACAATCAAC 643
 QY 644 CGGACACACCCCAAGGCTGGCTGGGGAAACCTTACCTCTCTGAGATTTTCATCATC 82647
 Db 82648 CGGACACACCCCAAGGCTGGCTGGGGAAACCTTACCTCTCTGAGATTTTCATCATC 82647
 QY 704 TCAAGTTCTCTTATATCCAGAGCAAGCAAGCAGATCATATATTAATTTATGACTTTATA 82707
 Db 82708 TCAAGTTCTCTTATATCCAGAGCAAGCAAGCAGATCATATATTAATTTATGACTTTATA 82707
 RESULT 4
 AC007193/c
 LOCUS: 146180 bp DNA linear PRI 17-JUN-1999
 DEFINITION AC007193 Homo sapiens chromosome 19, BAC 82621 (CIT-B-139a18), complete
 AC007193
 VERSION AC007193.1 GI:4558635
 KEYWORDS hmg.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 AUTHORS 1 (bases 1 to 146180)
 Lamerdin,J.E., McCready,P.M., Skowronski,E., Viswanathan,V.,
 Burkhart-Schultz,K.J., Gordon,L., Dias,J., Ramirez,M.,
 Stillwagen,S., Phan,H., Velasco,N., Do,L., Regala,W., Terry,A.,
 Avila,J., Dangnanan,L., Eiler,A., Christensen,M., Georgescu,A.,
 Amico-Keller,G., Coefield,J., Duarte,S., Lucas,S., Bruce,R.,
 Nolan,M., Truong,S., Kobayashi,A., Ariellano,A., Sanders,C., Ow,D.,
 Digs412
 Sequence analysis of a 1.9 Mb region in 19q13.2 between APOE and
 unpublished
 2 (bases 1 to 146180)
 Lamerdin,J.E.
 Direct Submission
 Submitted (02-APR-1999) Joint Genome Institute, Lawrence Livermore
 National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
 Map and sequence oriented from q centromere to telomere. BAC 82621
 (CIT-B-139a18) overlaps BAC 264576 (CIT-B-297n14) to the left from
 bases 1 to 27,487 of this accession, and separated from BAC 89981
 (CIT-B-156d10) on the right by a gap of approximately 17 to 23 kb.
 Additional chromosome 19 map and sequence information may be
 obtained at: <http://www-bio.lnl.gov/bdnp/genome/genome.html>.
 Location/Qualifiers
 1. 146180

```

repeat_region      1.101
/rpt_family="L1"
511.677
/rpt_family="L1"
1018.1158
/rpt_family="L1"
1159.1453
/rpt_family="AluX"
1459.1525
/rpt_family="L1"
complement(1857, .1974)
/rpt_family="FLAN_A"
2115.2375
/rpt_family="L1MB4"

```

```
repeat_region 2478..2549  
/rpt_family="LINE2"  
complement(2629..2921)  
repeat_region /rpt_family="AluJo"  
3355..3655  
repeat_region /rpt_family="AluSg"  
3673..3939  
/rpt_family="L1"  
repeat_region complement(4005..4264)  
/rpt_family="AluSx"  
complement(4269..4568)  
repeat_region /rpt_family="AluY"  
5064..5122  
/note="predicted exon, program: grail2exons_human_1.3,  
frame: 2, quality: excellent, score: 81.000"  
misc_feature  
/rpt_family="AluY"  
5626..5886  
/note="predicted exon, program: grail2exons_human_1.3,  
frame: 0, quality: good, score: 52.000"  
misc_feature  
7398..7444  
/rpt_family="Alu"  
7480..7595  
/rpt_family="(TA)n"  
7596..7812  
repeat_region /rpt_family="AluJb"  
7900..8191  
/rpt_family="AluSx"  
complement(8575..8597)  
repeat_region /rpt_family="AT-rich"  
complement(8598..8878)  
/rpt_family="AluSx"  
9477..9766  
repeat_region /rpt_family="AluSx"  
9997..10258  
/rpt_family="AluSg"  
10273..10347  
/rpt_family="FLAM"  
10691..10818  
repeat_region /rpt_family="AluJb"  
complement(10999..11283)  
repeat_region /rpt_family="AluJo"  
complement(11302..11517)  
repeat_region /rpt_family="L1MC/D"  
complement(11536..11661)  
repeat_region /rpt_family="FLAM/C"  
complement(11667..11798)  
repeat_region /rpt_family="L1MC/D"  
complement(11833..12138)  
repeat_region /rpt_family="AluSg"  
complement(12158..12478)  
repeat_region /rpt_family="L1MC/D"  
12479..12642  
/rpt_family="FRAM"  
complement(12655..12795)  
repeat_region /rpt_family="L1MC/D"  
complement(12806..13106)  
repeat_region /rpt_family="AluSx"  
complement(13356..13481)  
repeat_region /rpt_family="(GAAA)n"  
complement(13482..13694)  
repeat_region /rpt_family="AluJb"  
complement(13709..14037)  
repeat_region /rpt_family="L1"  
complement(14009..14105)  
repeat_region /rpt_family="L1MC/D"  
complement(14115..14415)  
repeat_region /rpt_family="AluY"  
complement(14417..14849)  
repeat_region /rpt_family="L1MC/D"  
complement(14859..15153)  
repeat_region /rpt_family="AluSx"  
15226..15492  
repeat_region
```

```
misc_feature  
/rpt_family="AluJo"  
complement(16257..16340)  
/note="predicted exon, program: grail2exons_human_1.3,  
frame: 0, quality: good, score: 63.000"  
repeat_region /rpt_family="AluSg"  
16362..16666  
repeat_region /rpt_family="AluSg"  
complement(16691..17060)  
repeat_region /rpt_family="THERIC"  
17482..17781  
/rpt_family="AluJo"  
17782..17870  
/rpt_family="(GAAA)n"  
17968..18273  
/rpt_family="AluSg"  
complement(18401..18817)  
repeat_region /rpt_family="L1R7"  
complement(18818..19193)  
repeat_region /rpt_family="HERVH"  
complement(19194..19412)  
/rpt_family="HERVH"  
complement(19414..20028)  
/rpt_family="HERVH"  
complement(20031..23647)  
repeat_region /rpt_family="HERVH"  
complement(23651..24066)  
/rpt_family="L1R7"  
complement(24403..24714)  
/rpt_family="HER74"  
24789..25080  
repeat_region /rpt_family="AluSx"  
25161..25207  
/rpt_family="L1MD3"  
25224..25509  
repeat_region /rpt_family="AluJb"  
25605..25901  
/rpt_family="AluSp"  
25902..26069  
/rpt_family="AluSg/x"  
complement(26224..26256)  
repeat_region /rpt_family="(TAA)n"  
complement(26257..26545)  
/rpt_family="AluJo"  
26573..26876  
repeat_region /rpt_family="AluJo"  
complement(27125..27422)  
repeat_region /rpt_family="AluSx"  
27866..28173  
repeat_region /rpt_family="AluSx"  
28181..28279  
/rpt_family="AluJ/FLAM"  
complement(28359..28522)  
repeat_region /rpt_family="AluSg"  
complement(28526..28705)  
repeat_region /rpt_family="AluSc"  
complement(28733..28872)  
repeat_region /rpt_family="AluJb"  
28874..29081  
/rpt_family="AluJo"  
29463..29583  
misc_feature  
/note="predicted exon, program: grail2exons_human_1.3,  
frame: 2, quality: good, score: 71.000"  
repeat_region complement(29585..29888)  
/rpt_family="AluSg"  
30336..30634  
/rpt_family="AluSg"  
30646..30779  
/rpt_family="AluJb"  
complement(30830..31040)  
repeat_region /rpt_family="MIR"
```

Query Match 18.48; Score 140.6; DB 9; Length 146180;
Best Local Similarity 78.08; Pred. No. 9.3e-28;
Matches 202; Conservative 0; Mismatches 39; Indels 18; Gaps 2;

13 GAGCCATAGGGCTCCCCGAGGGGTGCATGCTAGCTGTCTTTGGCATTTTCTGCATCTCCAGGC 78
Db 21 GAAGCATAGAGGCCACATATCGATCTTGGCTTTGTCTGCGAGTAACAGATCTTCCTCC 111
Qy 79 TCCCTGCTCACACGAGGCCCGCAGTGGCCCCATGACTCTCTTACCTGTAAGCTGCACAGC 80
Db 81 TCCAGTGTCTAAAGAAAGAACTACGAGACGCTCTGTTGGCTCAGAGACTGTGGCTGTGCCAGC 138
Qy 139 CACACAGAGATGTGGGGACAAGTTCTTACGACCCCCCTGACACTGTGCTATGATGATG 140
Db 141 CGACACCCAGGTGTGGGACAAGATCTTACAAACCTTGAAGACATGTGTTATATGATG 198
Qy 199 CCGTGGGCGCCTTGGCCAGGACGAGACGTTG---AAATGCACCTTGTAGAGTGTCT 200
Db 201 CCATCTTATCTCTTAAAGAGAGACCCGCCGCTGTGTGCTACACTGCACCTTCTGGCCCTGCT 255
Qy 256 TTAGAGAGTGTCCCTCCCTGAGCTT 280
Db 261 TTGAGCTGTGCTGTCCGAGTCTTT 285

RESULT 7					
AX118905					
LOCUS	AX118905				
DEFINITION	Sequence 69 from Patent WO0129221.	378 bp	DNA	linear	PAT 11-MAY-2001
ACCESSION	AX118905				
VERSION	AX118905.1				
KEYWORDS	GI:14035859				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 378)				
AUTHORS	Conklin,D.C. and Yee,D.P.				
TITLE	Proteins and polynucleotides encoding them				
JOURNAL	Patent: WO 0129221-A 69 26-APR-2001;				
	Zymogenetics, Inc. (US)				
FEATURES	Location/Qualifiers				
source	1..378				

```
CDS
    /db_xref="taxon:9606"
    1..378
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAC38526.1"
```



```

misc_feature
/rpt_family="MER64"
4294..4451
/Note="DSS similarity to overlapping ESTs:
N39722 xy92e04.r1 Homo sapiens cDNA clone 269214 5';
(256..438); 81% identity.-AA283147 zt14407.r1
NCI_CGAP_GCB1 Homo sapiens cDNA
clone IMAGE:713101 5'; (248..400); 99% identity.-AA1316
clone 505261 5'; (127..281); 97% identity.-N28707
xy66d11.r1 Homo sapiens cDNA clone 266709 5'; (67..222);
NbHPU Homo sapiens cDNA clone 505261 3'; (437..275); 97%
identity."
complement(4789..5083)
/rpt_family="AluSx"
5875..5898
/Note="DSS similarity to N28707 xy66d11.r1 Homo sapiens
cDNA clone 266709 5'; (223..246); 100% identity."
5875..5972
/Note="DSS similarity to overlapping ESTs:
N39722 xy92e04.r1 Homo sapiens cDNA clone 269214 5';
pregnant uterus NbHPU Homo sapiens cDNA clone 505261 5';
(282..379); 100% identity.-AA142861 z149d07.s1 Soares
pregnant uterus NbHPU Homo sapiens cDNA clone 505261 3';
(274..175); 95% identity."
5906..6036
/Note="Predicted exon, program: graal2exons_human_1.3,
frame: 1, quality: excellent, score: 100.000"
complement(6148..6443)
/rpt_family="AluSx"
complement(6608..6648)
/rpt_family="LINE2"
complement(7118..7340)
/rpt_family="MIR"
complement(7538..7639)
/rpt_family="(CA)n"
complement(7792..8054)
/rpt_family="MIR"
complement(8789..9083)
/rpt_family="AluG"
9456..9520
/rpt_family="MER5B"
complement(9539..9841)
/rpt_family="AluSx"
9909..9944
/rpt_family="MER5B"
10099..10401
/rpt_family="AluSx"
complement(10587..10752)
/rpt_family="AluJ"
complement(10753..11057)
/rpt_family="AluSx"
complement(11065..11191)
/rpt_family="AluJ"
11322..11619
/rpt_family="AluSg"
11775..11814
/rpt_family="AT-rich"
complement(11864..11885)
/rpt_family="AT-rich"
complement(11914..11976)
/rpt_family="AluS"
complement(12035..12068)
/rpt_family="AT-rich"
complement(12186..12327)
/rpt_family="L1M1"
complement(12334..12384)
/rpt_family="LINE2"
12784..12994
/Note="Predicted exon, program: graal2exons_human_1.3,
frame: 1, quality: excellent, score: 92.000"
12847..12908

```

```

misc.feature
12895      12894
/nrpt_family="GC_rich".
/note="DDBS similarity to AA765166 n279c05.s1 NCI_CGAP_GCB1
Homo sapiens cDNA clone IMAGE:1301672; (122, 123); 978
Identity."
13002, 13032
/rpt_family="GC_rich"
13080, 13162
/note="predicted exon, program: grail2exons_human_1.3,
frame: 2, quality: excellent, score:100.000_DDS
similarity to AA765166 n279c05.s1 NCI_CGAP_GCB1 Homo
sapiens cDNA clone IMAGE:1301672; (122, 48); 1008
complement(13759, 13954)
/rpt_family="GC_rich"

repeat_region
misc.feature

Query Match
Best Local Similarity 9.2%; Score 70.2; DB 9;
Matches 121; Conservative 0; Mismatches 63; Indels 3; Gaps 1;

0Y 97 CCCCAGTGGCCCCCATGATGCTCTTACCGATGCTGTGTGCCACACACAGAGATGTGGG 156
    |||||
Db 121896 CCCCAGACGCTCTGTGTGGCTCAGAGCTGTGGCTGCCACACACAGAGATGTGGG 156
    |||||

0Y 157 ACAAGTCTCAGACCCCTCAGACATGTTGCTATGATGATGCCGTGTCCTTGCCCA 121837
    |||||
Db 121836 ACAAGATCTCAACACCTTCAGACAGCAGTGTATGATGATGCCATCTTATCCTTAAG 121777
    |||||

0Y 217 GGACCCAGACGTTGG---AAATGCACCTTCAGAGTCTGCTTTGAGACATGCTGCCCT 273
    |||||
Db 121776 AATACCCGCGCTGTGGCTCAGCTGACCTTGTGGCCCTGCTTGAAGCTCTGCTGCCG 121717
    |||||

0Y 274 GGACCTT 280
    |||||
Db 121716 AGTCTTT 121710

```

[illegible]

```

source
1. .166500
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.3 between DM and c5R1"
/clone="CIT-B-344H19"
/clone_11d="Cal Tech CIT-B BAC library"
/notes="T1NL clone name: BC282485"
70. .145
/rpt_family="(TGA)n"
complement(177. .453)
/rpt_family="AluJb"
complement(472. .712)
/rpt_family="MIR"
746. .904
/misc_feature
/notes="DDS similarity to overlapping ESTs:
N39722 yx92e04.r1 Homo sapiens cDNA clone 269214 5';
(12. .170); 99% identity. -AA283147 z149d07.r1 NCI CGAP GCBI
Homo sapiens cDNA clone IMAGE:713101 5'; (6. .163); 99%
identity. -AA143160 z149d07.r1 Soares pregnant uterus Nbhpu
Homo sapiens cDNA clone 505261 5'; (1. .43); 100%
identity. -AA142881 z149d07.s1 Soares pregnant uterus Nbhpu
Homo sapiens cDNA clone 505261 3'; (570. .527); 64%
identity."
join<746. .904,1157. .1239,4294. .4451,5875. .6036,
12784. .12994,13080. .13162,13903. .13958,14053. .14128,
14429. .14568,14726. .14804,14982. .15039,15813. .15885,
16023. .16122)
/notes="Hypothetical arginine-rich gene product;
Hypothetical CDS constructed from overlapping ESTs and
Xiral predictions. EST coverage is lacking for some
portions of the CDS. Gene may be alternatively spliced, as
multiple transcripts map into this region; alternatively,
these individual clones may represent artifactual events
during RT and cDNA creation"
/codon_start=3
/evidence=not_experimental
/product="BC282485.1"
/protein_id="AA038244.1"
/db_xref="GI:5042405"
/translation="TLMADPAGQVDYVPRGVHAYVMVSGVLEVEDRMTADQ
MRGFEDAGFTEDLHTKGNKQNTFCFHMESATQSSSYTLDTLTDTDESLENK
MGGRPSIAPRSQNLNSKRYLILYVEFDRIHPLPYGKRPVYLGIIITSLKE
ELGLGLDQNTDRTRENEIMHLEQVSRILASCRLEADQDGSREKALGRAROE
AEALRGVRCLELEIROERGLGHVAGRGDCRLAELEAKERSLRLKLTIT
SEALYKRGKSGRGRPARSPSPSGRALRDPRTFAVAKERKREIOMKQORN
RLGGSGSDGPSVSMROTQPPALITGEGDAPNRSRNSSVDSRSCSSASCSDL
EDFSESLSRGCHRRGRKPPSPWGSNMKSPVRSRHOKSLANSQGWVPIKEYSSE
HOADMEDIDARKALOEYMNRLDMS"
1157. .1239
/notes="predicted exon, program: graal2exons_human_1.3,
frame: 0, quality: good, score: 62.000-DDS similarity to
overlapping ESTs:
N39722 yx92e04.r1 Homo sapiens cDNA clone 269214 5';
(171. .255); 95% identity. -AA283147 z149d07.r1
NCI CGAP GCBI Homo sapiens cDNA
clone IMAGE:713101 5'; (164. .247); 92% identity. -AA143160
z149d07.r1 Soares pregnant uterus Nbhpu Homo sapiens cDNA
clone 505261 5'; (44. .126); 100% identity. -N28707
yx66d11.r1 Homo sapiens cDNA clone 266709 5'; (1. .66); 98%
identity. -AA142881 z149d07.s1 Soares pregnant uterus Nbhpu
Homo sapiens cDNA clone 505261 3'; (526. .438); 82%
identity."
1536. .1634
/rpt_family="(CA)n"
complement(1673. .1973)
/rpt_family="AluSx"
complement(2040. .2185)
/rpt_family="LINE2"
complement(2200. .2498)
/rpt_family="AluSx"
complement(2576. .2746)
/rpt_family="AluSg/x"

repeat_region
complement(2747. .3044)
/rpt_family="AluSx"
3048. .3750
/rpt_family="MER64"
repeat_region
3841. .4060
/rpt_family="MER64"
4294. .4451
/misc_feature
/notes="DDS similarity to overlapping ESTs:
N39722 yx92e04.r1 Homo sapiens cDNA clone 269214 5';
(256. .438); 81% identity. -AA283147 z149d07.r1
NCI CGAP GCBI Homo sapiens cDNA
clone IMAGE:713101 5'; (248. .400); 99% identity. -AA143160
z149d07.r1 Soares pregnant uterus Nbhpu Homo sapiens cDNA
clone 505261 5'; (127. .281); 97% identity. -N28707
clone 505261 5'; (127. .281); 97% identity. -N28707
yx66d11.r1 Homo sapiens cDNA clone 266709 5'; (67. .222);
98% identity. -AA142881 z149d07.s1 Soares pregnant uterus
Nbhpu Homo sapiens cDNA clone 505261 3'; (437. .275); 97%
identity."
complement(4789. .5083)
/rpt_family="AluSx"
5875. .5898
/misc_feature
/notes="DDS similarity to N28707 yx66d11.r1 Homo sapiens
cDNA clone 266709 5'; (223. .246); 100% identity."
5875. .5972
/misc_feature
/notes="DDS similarity to overlapping ESTs:
N39722 yx92e04.r1 Homo sapiens cDNA clone 269214 5';
(439. .506); 71% identity. -AA143160 z149d07.r1 Soares
pregnant uterus Nbhpu Homo sapiens cDNA clone 505261 5';
(282. .379); 100% identity. -AA142881 z149d07.s1 Soares
pregnant uterus Nbhpu Homo sapiens cDNA clone 505261 3';
(274. .175); 95% identity."
5906. .6036
/misc_feature
/notes="predicted exon, program: graal2exons_human_1.3,
frame: 1, quality: excellent, score: 100.000"
complement(6148. .6443)
/rpt_family="AluSx"
complement(6608. .6648)
/rpt_family="LINE2"
complement(7118. .7340)
/rpt_family="MIR"
complement(7538. .7639)
/rpt_family="(CA)n"
complement(7792. .8054)
/rpt_family="MIR"
complement(8789. .9083)
/rpt_family="AluJc"
9456. .9520
/rpt_family="MER5B"
complement(9539. .9841)
/rpt_family="AluSx"
9909. .9944
/rpt_family="MER5B"
10099. .10401
/rpt_family="AluSx"
complement(10587. .10752)
/rpt_family="AluJ"
complement(10735. .11057)
/rpt_family="AluSx"
complement(11065. .11191)
/rpt_family="AluJ"
11322. .11619
/rpt_family="AluSg"
11775. .11814
/rpt_family="AT_rich"
complement(11864. .11885)
/rpt_family="AT_rich"
complement(11914. .11976)
/rpt_family="AluS"
complement(12053. .12088)
/rpt_family="AT_rich"
complement(12186. .12327)
/rpt_family="L1M1"
complement(12334. .12384)

```

```

misc_feature
12784..12994
/note="Predicted exon, program: grail2exons_human_1.3,
frame: 1, quality: excellent, score: 92.000"
repeat_region
12847..12808
/rpt_family="GC_rich"
misc_feature
12895..12994
/note="BDS similarity to AA765166 n279c05.s1 NCI_CGAP_GCB1
identity."
repeat_region
13002..13032
/rpt_family="GC_rich"
misc_feature
13080..13162
/note="Predicted exon, program: grail2exons_human_1.3,
frame: 2, quality: excellent, score: 100.000-BDS
similarity to AA765166 n279c05.s1 NCI_CGAP_GCB1 Homo
identity."
repeat_region
13799..13954
/rpt_family="GC_rich"

Query Match
Best Local Similarity 8.5%; Score 64.8; DB 9; Length 166500;
Matches 106; Conservative 0; Mismatches 47; Indels 3; Gaps 1;

QY 128 GCTGTGCGCCACACACAGATGTGGGACATTTCTACAGCCCTGACGACTGTG 187
|||||
Db 158370 GCTTCCACGCGCCGACCGAGGTGTGAGACACAGATCTACACCCCTTGAGCAGTCTG 187
|||||

QY 188 CTATGATGATGCGCTGTGCTGCTGCTGCGCAGACCCAGAGCTGTGG--AACTGACACTT 244
|||||
Db 158430 TTACATGATGCGCTGTGCTGCTGCTGCGCAGACCCAGATGTGTGCTGCTGCACTT 158429
|||||

QY 245 CAGAGCTGCTTTGAGACAGTGTGCGCCCTGACACTT 280
|||||
Db 158490 CTGGCCCTCTGCTGAGCTGCTGCTGTGATTCCTT 158525
|||||

RESULT 10
AF429315 125020 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens junctionphillin 3 (JPH3) gene, partial cds.
DEFINITION AF429315
ACCESSION AF429315.1 GI:17646244
VERSION AF429315.1
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 125020)
Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll,Ashworth,R.G., Flesher,A., Stevenin,G., Brice,A.,
A repeat expansion in the gene encoding junctionphillin-3 is
associated with Huntington disease-like 2
Nat. Genet. 29 (4), 377-378 (2001)
JOURNAL 21583737
MEDLINE 11694876
PUBMED 21583737
REFERENCE 2 (bases 1 to 125020)
Holmes,S.E., Ingersoll,Ashworth,R.G., Ross,C.A. and Margolis,R.L.
Direct Submission
Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
FEATURES
source
1. 125020
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q24.3; between D16S520 and W1-12410"
/note="Isolated from a patient with Huntington's
disease-Like 2 (HDL2)."
complement(35581..35746)
repeat_region
/rpt_type="tandem"

```

```

mRNA
complement(<36507..>36887)
gene
/product="junctionphillin 3"
/complement(<36507..>36887)
/feature="JPH3"
/note="JPH3"
complement(<36507..36887)
CDS
/gene="JPH3"
/note="JPH3"
/note="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/product="junctionphillin 3"
/protein_id="AAL40941.1"
/db_xref="GI:17646245"
/translation="MSGGRFEDGGSTCGGWDGKAHGVCTGPGGCGTGSWS
HGFEVLGYVTWPSGNTYGTGTAQGRHRHIGLESKGVYKGVTHGFGRTGVRECAQ
NGAYEGTWSNGIDGDTETYSDD"

BASE COUNT 29056 a 32731 c 30696 g 28283 t 4254 others
ORIGIN

Query Match
Best Local Similarity 6.2%; Score 47.2; DB 9; Length 125020;
Matches 75; Conservative 281; Mismatches 298; Indels 3; Gaps 2;

QY 1 CTCACCTGCAACACCCAGACCCCTGCTCCCGAGGCTGATGCTGTCTTGTCCA 60
|||||
Db 17022 CCKCMSSYTWMSASYSMSRRYWRKMSKWCMSRSKSSCKSRGCGSMGKGKSYG 17081
|||||

QY 61 TTTTCTGATCTCCAGGCTCTCTGCTGCTACACGAGACCCCA--GTGGCCCTATCTCC 118
|||||
Db 17082 RTKRKRSRKMGMKAMKMYMYRSMRMKMYSSKGMTCMYCWCGRGRTCTCSMTSRAMC 17141
|||||

QY 119 TTACCGATGCTGTGCGCCACACACAGAGATGTGGGCAAGTCTTACAGCCCTGCA 178
|||||
Db 17142 GSYCTAKCKSMSCYSCYVGMKGYIYACYSKSSMSKCYMRSTYSTGCGCCCTTTTTC 17201
|||||

QY 179 GCACGTGTCTATGATGATGCCGCTGCTGCTGCGCCAGACCCAGAGCTGTGAAGTCC 17201
|||||
Db 17202 CCCNANTGGGAACTTTTNCNKRTYSYRRKNCAMCKYNNYNSVRSRAGMSTCK 17261
|||||

QY 239 CACCTTGAAGTCTGCTTTGAGCAGTGTGCTGCTGCTGACCTTCATGGAAGTGAATAA 298
|||||
Db 17262 YKSSMTMSMASYCMCMNMYICSMMSMASRGMSSSYKMKMSRMSKCYMKCMKMCY 17321
|||||

QY 299 C-CAGAACTGCGACTAGCCCGGACCTGGATGACAGAGCTTTGTGCAGTGTCAATAA 357
|||||
Db 17322 CMKRSRSMGMSYMYAASWKSSSRGCTRCIWCMSKCYKSYMMMSKRYKMKWSR 17381
|||||

QY 358 GGAACATCAGGGAAGAGATCTCTGATTTCTCTCTGCTGGTGGGCTTGAGAAAGAG 417
|||||
Db 17382 CWSMWSASBSKCYKYSKMSKMSKCYRCACCMKMGYMYRMCMWSMKRMMKKS 17441
|||||

QY 418 GCTGTGTACCTAGATCTGGATGCTGAGTGTGCTGTTGGGGCCAGAGAAACACACA 477
|||||
Db 17442 AMYRRMRMRMGAMMMCMKCYRMSRCMMKYSYCASCRSCAMMSGYMKACAYMMSS 17501
|||||

QY 478 CTCAACTGCCCACTTCTTGTGACCTGTCTGAGGCCACCCGCGCTGCTGAGGA 537
|||||
Db 17502 YWCTCMTSTYSYRCCMCCKMSYKSKSSMSYKSKSSMSYKGRKRSYMCCTRGANSCW 17561
|||||

QY 538 GGGCCACAGGTCCCTTCTAGAAATTCTGACACATAGATGTGCTGTGATGGGGCC 597
|||||
Db 17562 RRCYWRGASMSRRAGMSRRRAKGRSMGGRSKWMYMGMRKRYIYCTGRBMMTYMCC 17621
|||||

QY 598 CAGGAGACTGAACTCTCTGATGAGCCCTATGCCAACATACCCGACACACC 654
|||||
Db 17622 CMRRSYMYRSMAMGRKSSMSGRMMGASMSRRCKSASRSSCSRRMKRMSRC 17678
|||||

RESULT 11
AC013622_0/c
WPCOMMENT

```

Sequence split into 5 fragments Locus AC013622 Accession AC013622

Fragment Name Begin End
AC013622_0 1 110000
AC013622_1 100001 210000
AC013622_2 200001 310000
AC013622_3 300001 410000
AC013622_4 400001 436313
AC013622 436313 bp DNA linear HTG 30-MAY-2000
Mus musculus clone RP23-232H18, *** SEQUENCING IN PROGRESS ***, 52
EF171717
Mus musculus clone RP23-232H18, *** SEQUENCING IN PROGRESS ***, 52
unpublished pieces.

CCESION AC013622
AC013622.4 GI:8102534
HTG: HTGS.PHASE1.
HOUSE MOUSE

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 436313)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
Unpublished
2 (bases 1 to 436313)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckert, R., Boguslavsky, L., Bouckgalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., Dattalano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
Galan, J., Gardy, S., Grant, G., Hagos, B., Harford, A., Horton, L.,
Howland, J., Johnson, R., Jones, C., Kahn, L., Karatas, A., Klein, J.,
Lehocky, J., Liu, C., Locke, R., Macdonald, P., Marquis, N.,
McEwan, P., McGuck, A., McKernan, K., McDonald, J., Meldrum, J.,
Merrill, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tessier, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (11-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 30, 2000 this sequence version replaced g1.7329381.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996:1997).
http://ftp.genome.washington.edu/RW/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIMR
Web site: http://www.seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
Project Information
Center project name: 232_H18

NOTE: This is a 'working draft' sequence. It currently
consists of 52 contigs. The true order of the pieces is
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 2045: contig of 2045 bp in length
2046 2145: gap of 100 bp
2146 4521: contig of 2376 bp in length
4522 4621: gap of 100 bp
4622 8024: contig of 3403 bp in length
8025 8124: gap of 100 bp
8125 11099: contig of 2975 bp in length
11100 11199: gap of 100 bp
11200 13868: contig of 2669 bp in length
13869 13968: gap of 100 bp
13969 18430: contig of 4462 bp in length
18431 18530: gap of 100 bp
18531 21696: contig of 3166 bp in length

21697 21796: gap of 100 bp
21797 25742: contig of 25742 bp in length
25742 25841: gap of 100 bp
25842 29357: contig of 29357 bp in length
29358 29457: gap of 100 bp
29458 32719: contig of 32719 bp in length
32720 32819: gap of 100 bp
32820 35995: contig of 35995 bp in length
35996 36095: gap of 100 bp
36096 41396: contig of 538 bp in length
41397 41496: gap of 100 bp
41497 45153: contig of 3657 bp in length
45154 45253: gap of 100 bp
45254 49318: contig of 4065 bp in length
49319 49418: gap of 100 bp
49419 54714: contig of 5296 bp in length
54715 54814: gap of 100 bp
54815 58285: contig of 3471 bp in length
58286 58385: gap of 100 bp
58386 63447: contig of 5062 bp in length
63448 63547: gap of 100 bp
63548 67966: contig of 4419 bp in length
67967 68066: gap of 100 bp
68067 73138: contig of 5072 bp in length
73139 73238: gap of 100 bp
73239 77301: contig of 4063 bp in length
77302 77401: gap of 100 bp
77402 82081: contig of 4680 bp in length
82082 82181: gap of 100 bp
82182 86119: contig of 4538 bp in length
86120 86819: gap of 100 bp
86820 92113: contig of 5294 bp in length
92114 92213: gap of 100 bp
92214 96761: contig of 4548 bp in length
96762 96861: gap of 100 bp
96862 101458: contig of 4597 bp in length
101459 101558: gap of 100 bp
101559 107167: contig of 5609 bp in length
107168 107267: gap of 100 bp
107268 113324: contig of 6057 bp in length
113325 113424: gap of 100 bp
113425 119681: contig of 6257 bp in length
119682 119781: gap of 100 bp
119782 126056: contig of 6275 bp in length
126057 126156: gap of 100 bp
126157 132230: contig of 6074 bp in length
132231 132330: gap of 100 bp
132332 140500: contig of 8170 bp in length
140501 140600: gap of 100 bp
140601 149110: contig of 8510 bp in length
149111 149210: gap of 100 bp
149211 156954: contig of 7744 bp in length
156955 157054: gap of 100 bp
157055 163961: contig of 6907 bp in length
163962 164061: gap of 100 bp
164062 172377: contig of 8316 bp in length
172378 172477: gap of 100 bp
172478 181342: contig of 8865 bp in length
181343 181442: gap of 100 bp
181443 189265: contig of 7823 bp in length
189266 189365: gap of 100 bp
189366 197225: contig of 7860 bp in length
197226 197325: gap of 100 bp
197326 206982: contig of 9657 bp in length
206983 207082: gap of 100 bp
207083 215973: contig of 8891 bp in length
215974 216073: gap of 100 bp
216074 227424: contig of 11351 bp in length
227425 227524: gap of 100 bp
227525 237208: contig of 9684 bp in length
237209 237308: gap of 100 bp
237309 247157: contig of 9849 bp in length
247158 247257: gap of 100 bp

21696-21796 bp in length
25742-25841 bp in length
29357-29457 bp in length
32719-32819 bp in length
35995-36095 bp in length
41396-41496 bp in length
45153-45253 bp in length
49318-49418 bp in length
54714-54814 bp in length
58285-58385 bp in length
63447-63547 bp in length
67966-68066 bp in length
73138-73238 bp in length
77301-77401 bp in length
82081-82181 bp in length
86119-86819 bp in length
92113-92213 bp in length
96761-96861 bp in length
101458-101558 bp in length
107167-107267 bp in length
113324-113424 bp in length
119681-119781 bp in length
126056-126156 bp in length
132230-132330 bp in length
140500-140600 bp in length
149110-149210 bp in length
156954-157054 bp in length
163961-164061 bp in length
172377-172477 bp in length
181342-181442 bp in length
189265-189365 bp in length
197225-197325 bp in length
206982-207082 bp in length
215973-216073 bp in length
227424-227524 bp in length
237208-237308 bp in length
247157-247257 bp in length

247258 261074: contig of 13817 bp in length
 * 261075 261174: gap of 100 bp
 * 261175 274135: contig of 12961 bp in length
 * 274136 274235: gap of 100 bp
 * 274236 289816: contig of 15581 bp in length
 * 289817 289916: gap of 100 bp
 * 289917 305346: contig of 15430 bp in length
 * 305347 305446: gap of 100 bp
 * 305447 320002: contig of 14556 bp in length
 * 320003 320102: gap of 100 bp
 * 320103 339946: contig of 18844 bp in length
 * 339947 340046: gap of 100 bp
 * 340047 366337: contig of 26591 bp in length
 * 366338 366737: gap of 100 bp
 * 366738 396109: contig of 29372 bp in length
 * 396110 396209: gap of 100 bp
 * 396210 436313: contig of 40104 bp in length.
 Location/Qualifiers
 1. 436313
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_lib="RP23-233H18"
 /clone_id="RP23-233H18"

BASE COUNT 123368 a 96120 c 95329 g 116149 t 6347 others
 ORIGIN

Query Match
 Best Local Similarity 5.3%; Score 40.8; DB 2; Length 110000;
 Matches 69; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 426 TACCTGATCTGAGATCTGAGTCTGTTGGGGCCAGAGAACACACTCACTG 485
 Db 109402 TTCTGCTGCGACGAGAGTCTGGGTGGGTCAGTACACTTCACTCACTCACTT 109343
 Qy 486 CCCACTTCTCTGACTCTGAGGCCACCTGCCCTGCTGAGGAGGCC 541
 Db 109342 CCACCTTCTACTCCACATGAATGGCGGAGAGACTCCACTCCCTGAGAGAGCAG 109287

RESULT 12
 AC013622_1/c
 WPCOMMENT
 Sequence split into 5 fragments LOCUS AC013622 Accession AC013622

Fragment Name Begin End
 AC013622_0 1 110000
 AC013622_1 100001 210000
 AC013622_2 200001 310000
 AC013622_3 300001 410000
 AC013622_4 400001 436313
 Continuation (2 of 5) of AC013622 from base 100001 (AC013622 Mus musculus clone RP23-233)

Query Match
 Best Local Similarity 5.3%; Score 40.8; DB 2; Length 110000;
 Matches 69; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 426 TACCTGATCTGAGATCTGAGTCTGTTGGGGCCAGAGAACACACTCACTG 485
 Db 9402 TTCTGCTGCGACGAGAGTCTGGGTGGGTCAGTACACTTCACTCACTCACTT 109343
 Qy 486 CCCACTTCTCTGACTCTGAGGCCACCTGCCCTGCTGAGGAGGCC 541
 Db 9342 CCACCTTCTACTCCACATGAATGGCGGAGAGACTCCACTCCCTGAGAGAGCAG 9287

RESULT 13
 AC021632_0/c
 WPCOMMENT
 Sequence split into 5 fragments LOCUS AC021632 Accession AC021632

Fragment Name Begin End
 AC021632_0 1 110000
 AC021632_1 100001 210000
 AC021632_2 200001 310000
 AC021632_3 300001 410000

LOCUS AC021632_4 400001 455450
 DEFINITION Mus musculus chromosome 11 clone MGS1-21363 map 11, *** SEQUENCING
 ACCESSION AC021632
 IN PROGRESS ***
 VERSION AC021632.4 GI:8389557
 KEYWORDS HTGS-PHASE1.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCES
 1 (bases 1 to 455450)
 Birren, B., Linton, L., Nusbaum, C., Allen, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Beckert, R., Bedalov, A.,
 Boguski, M.S., Boulikas, B., Brown, A., Burkett, G., Castelle, A.,
 Chao, Y., Collins, S., Collins, S., Collins, S., Collins, S., Collins, S.,
 DeFelice, K., Dewar, K., Dominko, M., Doyle, M., Fenech, P.,
 Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
 Garg, S., Grant, G., Hargis, B., Heaford, A., Horton, L.,
 Howland, J.C., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J.,
 Lander, T., Lech, J., Levine, R., Liew, C., Liu, G., Locke, K.,
 MacDonald, P., Margulis, N., McKernan, P., McGurk, A., McKernan, K.,
 McPherson, R., Meldrum, J., Menus, L., Morrow, J., Naylor, J.,
 Norman, C., O'Connor, T., O'Donnell, P., Olliver, T.M., Peterson, K.,
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange, B., Stange, B.,
 Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Tirrell, A., Vassiliou, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,
 Zimmer, A., and Zody, M.
 Direct Submission
 Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RV/RepeatMasker.html

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
 Web site: http://www.seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 Project Information
 Center project name: L5902
 Center clone name: 213_E_3

NOTE: This is a 'working draft' sequence. It currently
 consists of 56 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 7029: contig of 7029 bp in length
 7030 7129: gap of 100 bp
 7130 16474: contig of 9345 bp in length
 16475 16574: gap of 100 bp
 16575 25724: contig of 9150 bp in length
 25725 25824: gap of 100 bp
 25825 37070: contig of 11243 bp in length
 37071 37170: gap of 100 bp
 37171 47434: contig of 10667 bp in length
 47435 47534: gap of 100 bp
 47535 58615: contig of 11180 bp in length
 58615 58615: gap of 100 bp
 58616 58616: gap of 100 bp
 58617 58617: gap of 100 bp
 58618 58618: gap of 100 bp
 58619 58619: gap of 100 bp
 58620 58620: gap of 100 bp
 58621 58621: gap of 100 bp
 58622 58622: gap of 100 bp
 58623 58623: gap of 100 bp
 58624 58624: gap of 100 bp
 58625 58625: gap of 100 bp
 58626 58626: gap of 100 bp
 58627 58627: gap of 100 bp
 58628 58628: gap of 100 bp
 58629 58629: gap of 100 bp
 58630 58630: gap of 100 bp
 58631 58631: gap of 100 bp
 58632 58632: gap of 100 bp
 58633 58633: gap of 100 bp
 58634 58634: gap of 100 bp
 58635 58635: gap of 100 bp
 58636 58636: gap of 100 bp
 58637 58637: gap of 100 bp
 58638 58638: gap of 100 bp
 58639 58639: gap of 100 bp
 58640 58640: gap of 100 bp
 58641 58641: gap of 100 bp
 58642 58642: gap of 100 bp
 58643 58643: gap of 100 bp
 58644 58644: gap of 100 bp
 58645 58645: gap of 100 bp
 58646 58646: gap of 100 bp
 58647 58647: gap of 100 bp
 58648 58648: gap of 100 bp
 58649 58649: gap of 100 bp
 58650 58650: gap of 100 bp
 58651 58651: gap of 100 bp
 58652 58652: gap of 100 bp
 58653 58653: gap of 100 bp
 58654 58654: gap of 100 bp
 58655 58655: gap of 100 bp
 58656 58656: gap of 100 bp
 58657 58657: gap of 100 bp
 58658 58658: gap of 100 bp
 58659 58659: gap of 100 bp
 58660 58660: gap of 100 bp
 58661 58661: gap of 100 bp
 58662 58662: gap of 100 bp
 58663 58663: gap of 100 bp
 58664 58664: gap of 100 bp
 58665 58665: gap of 100 bp
 58666 58666: gap of 100 bp
 58667 58667: gap of 100 bp
 58668 58668: gap of 100 bp
 58669 58669: gap of 100 bp
 58670 58670: gap of 100 bp
 58671 58671: gap of 100 bp
 58672 58672: gap of 100 bp
 58673 58673: gap of 100 bp
 58674 58674: gap of 100 bp
 58675 58675: gap of 100 bp
 58676 58676: gap of 100 bp
 58677 58677: gap of 100 bp
 58678 58678: gap of 100 bp
 58679 58679: gap of 100 bp
 58680 58680: gap of 100 bp
 58681 58681: gap of 100 bp
 58682 58682: gap of 100 bp
 58683 58683: gap of 100 bp
 58684 58684: gap of 100 bp
 58685 58685: gap of 100 bp
 58686 58686: gap of 100 bp
 58687 58687: gap of 100 bp
 58688 58688: gap of 100 bp
 58689 58689: gap of 100 bp
 58690 58690: gap of 100 bp
 58691 58691: gap of 100 bp
 58692 58692: gap of 100 bp
 58693 58693: gap of 100 bp
 58694 58694: gap of 100 bp
 58695 58695: gap of 100 bp
 58696 58696: gap of 100 bp
 58697 58697: gap of 100 bp
 58698 58698: gap of 100 bp
 58699 58699: gap of 100 bp
 58700 58700: gap of 100 bp
 58701 58701: gap of 100 bp
 58702 58702: gap of 100 bp
 58703 58703: gap of 100 bp
 58704 58704: gap of 100 bp
 58705 58705: gap of 100 bp
 58706 58706: gap of 100 bp
 58707 58707: gap of 100 bp
 58708 58708: gap of 100 bp
 58709 58709: gap of 100 bp
 58710 58710: gap of 100 bp
 58711 58711: gap of 100 bp
 58712 58712: gap of 100 bp
 58713 58713: gap of 100 bp
 58714 58714: gap of 100 bp
 58715 58715: gap of 100 bp
 58716 58716: gap of 100 bp
 58717 58717: gap of 100 bp
 58718 58718: gap of 100 bp
 58719 58719: gap of 100 bp
 58720 58720: gap of 100 bp
 58721 58721: gap of 100 bp
 58722 58722: gap of 100 bp
 58723 58723: gap of 100 bp
 58724 58724: gap of 100 bp
 58725 58725: gap of 100 bp
 58726 58726: gap of 100 bp
 58727 58727: gap of 100 bp
 58728 58728: gap of 100 bp
 58729 58729: gap of 100 bp
 58730 58730: gap of 100 bp
 58731 58731: gap of 100 bp
 58732 58732: gap of 100 bp
 58733 58733: gap of 100 bp
 58734 58734: gap of 100 bp
 58735 58735: gap of 100 bp
 58736 58736: gap of 100 bp
 58737 58737: gap of 100 bp
 58738 58738: gap of 100 bp
 58739 58739: gap of 100 bp
 58740 58740: gap of 100 bp
 58741 58741: gap of 100 bp
 58742 58742: gap of 100 bp
 58743 58743: gap of 100 bp
 58744 58744: gap of 100 bp
 58745 58745: gap of 100 bp
 58746 58746: gap of 100 bp
 58747 58747: gap of 100 bp
 58748 58748: gap of 100 bp
 58749 58749: gap of 100 bp
 58750 58750: gap of 100 bp
 58751 58751: gap of 100 bp
 58752 58752: gap of 100 bp
 58753 58753: gap of 100 bp
 58754 58754: gap of 100 bp
 58755 58755: gap of 100 bp
 58756 58756: gap of 100 bp
 58757 58757: gap of 100 bp
 58758 58758: gap of 100 bp
 58759 58759: gap of 100 bp
 58760 58760: gap of 100 bp
 58761 58761: gap of 100 bp
 58762 58762: gap of 100 bp
 58763 58763: gap of 100 bp
 58764 58764: gap of 100 bp
 58765 58765: gap of 100 bp
 58766 58766: gap of 100 bp
 58767 58767: gap of 100 bp
 58768 58768: gap of 100 bp
 58769 58769: gap of 100 bp
 58770 58770: gap of 100 bp
 58771 58771: gap of 100 bp
 58772 58772: gap of 100 bp
 58773 58773: gap of 100 bp
 58774 58774: gap of 100 bp
 58775 58775: gap of 100 bp
 58776 58776: gap of 100 bp
 58777 58777: gap of 100 bp
 58778 58778: gap of 100 bp
 58779 58779: gap of 100 bp
 58780 58780: gap of 100 bp
 58781 58781: gap of 100 bp
 58782 58782: gap of 100 bp
 58783 58783: gap of 100 bp
 58784 58784: gap of 100 bp
 58785 58785: gap of 100 bp
 58786 58786: gap of 100 bp
 58787 58787: gap of 100 bp
 58788 58788: gap of 100 bp
 58789 58789: gap of 100 bp
 58790 58790: gap of 100 bp
 58791 58791: gap of 100 bp
 58792 58792: gap of 100 bp
 58793 58793: gap of 100 bp
 58794 58794: gap of 100 bp
 58795 58795: gap of 100 bp
 58796 58796: gap of 100 bp
 58797 58797: gap of 100 bp
 58798 58798: gap of 100 bp
 58799 58799: gap of 100 bp
 58800 58800: gap of 100 bp
 58801 58801: gap of 100 bp
 58802 58802: gap of 100 bp
 58803 58803: gap of 100 bp
 58804 58804: gap of 100 bp
 58805 58805: gap of 100 bp
 58806 58806: gap of 100 bp
 58807 58807: gap of 100 bp
 58808 58808: gap of 100 bp
 58809 58809: gap of 100 bp
 58810 58810: gap of 100 bp
 58811 58811: gap of 100 bp
 58812 58812: gap of 100 bp
 58813 58813: gap of 100 bp
 58814 58814: gap of 100 bp
 58815 58815: gap of 100 bp
 58816 58816: gap of 100 bp
 58817 58817: gap of 100 bp
 58818 58818: gap of 100 bp
 58819 58819: gap of 100 bp
 58820 58820: gap of 100 bp
 58821 58821: gap of 100 bp
 58822 58822: gap of 100 bp
 58823 58823: gap of 100 bp
 58824 58824: gap of 100 bp
 58825 58825: gap of 100 bp
 58826 58826: gap of 100 bp
 58827 58827: gap of 100 bp
 58828 58828: gap of 100 bp
 58829 58829: gap of 100 bp
 58830 58830: gap of 100 bp
 58831 58831: gap of 100 bp
 58832 58832: gap of 100 bp
 58833 58833: gap of 100 bp
 58834 58834: gap of 100 bp
 58835 58835: gap of 100 bp
 58836 58836: gap of 100 bp
 58837 58837: gap of 100 bp
 58838 58838: gap of 100 bp
 58839 58839: gap of 100 bp
 58840 58840: gap of 100 bp
 58841 58841: gap of 100 bp
 58842 58842: gap of 100 bp
 58843 58843: gap of 100 bp
 58844 58844: gap of 100 bp
 58845 58845: gap of 100 bp
 58846 58846: gap of 100 bp
 58847 58847: gap of 100 bp
 58848 58848: gap of 100 bp
 58849 58849: gap of 100 bp
 58850 58850: gap of 100 bp
 58851 58851: gap of 100 bp
 58852 58852: gap of 100 bp
 58853 58853: gap of 100 bp
 58854 58854: gap of 100 bp
 58855 58855: gap of 100 bp
 58856 58856: gap of 100 bp
 58857 58857: gap of 100 bp
 58858 58858: gap of 100 bp
 58859 58859: gap of 100 bp
 58860 58860: gap of 100 bp
 58861 58861: gap of 100 bp
 58862 58862: gap of 100 bp
 58863 58863: gap of 100 bp
 58864 58864: gap of 100 bp
 58865 58865: gap of 100 bp
 58866 58866: gap of 100 bp
 58867 58867: gap of 100 bp
 58868 58868: gap of 100 bp
 58869 58869: gap of 100 bp
 58870 58870: gap of 100 bp
 58871 58871: gap of 100 bp
 58872 58872: gap of 100 bp
 58873 58873: gap of 100 bp
 58874 58874: gap of 100 bp
 58875 58875: gap of 100 bp
 58876 58876: gap of 100 bp
 58877 58877: gap of 100 bp
 58878 58878: gap of 100 bp
 58879 58879: gap of 100 bp
 58880 58880: gap of 100 bp
 58881 58881: gap of 100 bp
 58882 58882: gap of 100 bp
 58883 58883: gap of 100 bp
 58884 58884: gap of 100 bp
 58885 58885: gap of 100 bp
 58886 58886: gap of 100 bp
 58887 58887: gap of 100 bp
 58888 58888: gap of 100 bp
 58889 58889: gap of 100 bp
 58890 58890: gap of 100 bp
 58891 58891: gap of 100 bp
 58892 58892: gap of 100 bp
 58893 58893: gap of 100 bp
 58894 58894: gap of 100 bp
 58895 58895: gap of 100 bp
 58896 58896: gap of 100 bp
 58897 58897: gap of 100 bp
 58898 58898: gap of 100 bp
 58899 58899: gap of 100 bp
 58900 58900: gap of 100 bp
 58901 58901: gap of 100 bp
 58902 58902: gap of 100 bp
 58903 58903: gap of 100 bp
 58904 58904: gap of 100 bp
 58905 58905: gap of 100 bp
 58906 58906: gap of 100 bp
 58907 58907: gap of 100 bp
 58908 58908: gap of 100 bp
 58909 58909: gap of 100 bp
 58910 58910: gap of 100 bp
 58911 58911: gap of 100 bp
 58912 58912: gap of 100 bp
 58913 58913: gap of 100 bp
 58914 58914: gap of 100 bp
 58915 58915: gap of 100 bp
 58916 58916: gap of 100 bp
 58917 58917: gap of 100 bp
 58918 58918: gap of 100 bp
 58919 58919: gap of 100 bp
 58920 58920: gap of 100 bp
 58921 58921: gap of 100 bp
 58922 58922: gap of 100 bp
 58923 58923: gap of 100 bp
 58924 58924: gap of 100 bp
 58925 58925: gap of 100 bp
 58926 58926: gap of 100 bp
 58927 58927: gap of 100 bp
 58928 58928: gap of 100 bp
 58929 58929: gap of 100 bp
 58930 58930: gap of 100 bp
 58931 58931: gap of 100 bp
 58932 58932: gap of 100 bp
 58933 58933: gap of 100 bp
 58934 58934: gap of 100 bp
 58935 58935: gap of 100 bp
 58936 58936: gap of 100 bp
 58937 58937: gap of 100 bp
 58938 58938: gap of 100 bp
 58939 58939: gap of 100 bp
 58940 58940: gap of 100 bp
 58941 58941: gap of 100 bp
 58942 58942: gap of 100 bp
 58943 58943: gap of 100 bp
 58944 58944: gap of 100 bp
 58945 58945: gap of 100 bp
 58946 58946: gap of 100 bp
 58947 58947: gap of 100 bp
 58948 58948: gap of 100 bp
 58949 58949: gap of 100 bp
 58950 58950: gap of 100 bp
 58951 58951: gap of 100 bp
 58952 58952: gap of 100 bp
 58953 58953: gap of 100 bp
 58954 58954: gap of 100 bp
 58955 58955: gap of 100 bp
 58956 58956: gap of 100 bp
 58957 58957: gap of 100 bp
 58958 58958: gap of 100 bp
 58959 58959: gap of 100 bp
 58960 58960: gap of 100 bp
 58961 58961: gap of 100 bp
 58962 58962: gap of 100 bp
 58963 58963: gap of 100 bp
 58964 58964: gap of 100 bp
 58965 58965: gap of 100 bp
 58966 58966: gap of 100 bp
 58967 58967: gap of 100 bp
 58968 58968: gap of 100 bp
 58969 58969: gap of 100 bp
 58970 58970: gap of 100 bp
 58971 58971: gap of 100 bp
 58972 58972: gap of 100 bp
 58973 58973: gap of 100 bp
 58974 58974: gap of 100 bp
 58975 58975: gap of 100 bp
 58976 58976: gap of 100 bp
 58977 58977: gap of 100 bp
 58978 58978: gap of 100 bp
 58979 58979: gap of 100 bp
 58980 58980: gap of 100 bp
 58981 58981: gap of 100 bp
 58982 58982: gap of 100 bp
 58983 58983: gap of 100 bp
 58984 58984: gap of 100 bp
 58985 58985: gap of 100 bp
 58986 58986: gap of 100 bp

103018	120081:	contlg	of	17064	bp	in	length
120082	120181:	gap	of	100	bp		
120182	136051:	contlg	of	15876	bp	in	length
136058	136157:	gap	of	100	bp		
136158	156302:	contlg	of	20145	bp	in	length
156303	156402:	gap	of	100	bp		
156403	167787:	contlg	of	11385	bp	in	length
167788	183109:	contlg	of	15222	bp	in	length
183110	183309:	gap	of	100	bp		
183210	196606:	contlg	of	13397	bp	in	length
196607	196706:	gap	of	100	bp		
196707	218924:	contlg	of	22218	bp	in	length
218925	219024:	gap	of	100	bp		
219025	249572:	contlg	of	30548	bp	in	length
249573	249672:	gap	of	100	bp		
249673	287186:	contlg	of	37516	bp	in	length
287189	287288:	gap	of	100	bp		
287289	289388:	contlg	of	2100	bp	in	length
289389	289488:	gap	of	100	bp		
289489	291859:	contlg	of	2371	bp	in	length
291860	291959:	gap	of	100	bp		
291960	294126:	contlg	of	2169	bp	in	length
294129	294228:	gap	of	100	bp		
294229	296882:	contlg	of	2624	bp	in	length
296883	296982:	gap	of	100	bp		
296983	299603:	contlg	of	2621	bp	in	length
299604	299703:	gap	of	100	bp		
299704	301793:	contlg	of	2030	bp	in	length
301794	301893:	gap	of	100	bp		
301894	305076:	contlg	of	3183	bp	in	length
305077	305176:	gap	of	100	bp		
305177	308174:	contlg	of	2998	bp	in	length
308175	308274:	gap	of	100	bp		
308275	310542:	contlg	of	2268	bp	in	length
310543	310642:	gap	of	100	bp		
310643	313870:	contlg	of	3228	bp	in	length
313871	313970:	gap	of	100	bp		
313971	317124:	contlg	of	3154	bp	in	length
317125	317224:	gap	of	100	bp		
317225	320595:	contlg	of	3371	bp	in	length
320596	320695:	gap	of	100	bp		
320696	323155:	contlg	of	2460	bp	in	length
323156	323255:	gap	of	100	bp		
323256	326284:	contlg	of	3029	bp	in	length
326285	326384:	gap	of	100	bp		
326385	329656:	contlg	of	3242	bp	in	length
329627	329726:	gap	of	100	bp		
329727	333482:	contlg	of	3756	bp	in	length
333483	333582:	gap	of	100	bp		
333583	333789:	contlg	of	3607	bp	in	length
337190	337289:	gap	of	100	bp		
337290	341875:	contlg	of	4586	bp	in	length
341876	341975:	gap	of	100	bp		
341976	344388:	contlg	of	2413	bp	in	length
344389	344488:	gap	of	100	bp		
344489	349154:	contlg	of	4666	bp	in	length
349155	349254:	gap	of	100	bp		
349255	353824:	contlg	of	4570	bp	in	length
353825	353924:	gap	of	100	bp		
353925	357953:	contlg	of	4029	bp	in	length
357954	358053:	gap	of	100	bp		
358054	363655:	contlg	of	5642	bp	in	length
363656	363795:	gap	of	100	bp		
363796	367278:	contlg	of	3483	bp	in	length
367279	367378:	gap	of	100	bp		

```

FEATURES
SOURCE
*
382172 382271: gap of 100 bp
*
382272 388633: contig of 6562 bp in length
*
388634 388933: gap of 100 bp
*
388934 393530: contig of 4537 bp in length
*
393531 393630: gap of 100 bp
*
393631 398934: contig of 5304 bp in length
*
398935 399034: gap of 100 bp
*
399035 404551: contig of 5517 bp in length
*
404552 404651: gap of 100 bp
*
404652 410493: contig of 5842 bp in length
*
410494 410593: gap of 100 bp
*
410594 415944: contig of 5351 bp in length
*
415945 416044: gap of 100 bp
*
416045 424416: contig of 8372 bp in length
*
424417 424516: gap of 100 bp
*
424517 430287: contig of 5771 bp in length
*
430288 430387: gap of 100 bp
*
430388 437707: contig of 7320 bp in length
*
437708 437807: gap of 100 bp
*
437808 445871: contig of 8064 bp in length
*
445872 445971: gap of 100 bp
*
445972 455450: contig of 9479 bp in length.
*
Location/Qualifiers
1..455450
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/map="11"
/clone="MGSI-213E3"
/clone_11b="Genome Systems Mouse BAC 1"
119379 a 104363 c 107639 g 117130 t 6939 others

```

```

Query Match          5.3%; Score 40.8; DB 2; Length 110000;
Best Local Similarity 59.5%; Pred. No. 2.1;
Matches 69; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

OY 426 TACCGAGATCMGGGATGCTGAGTGTGCTGTGTGGGGCCAGAGAAACACACTCAACTG 485
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 18530 TTCTCGCGCCAGGAGAGCTGGGTGGGTGCATGCTACACTTACACTACACTCAACTT 18471
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 486 CCCACTTCATTTGTGAGCTGTGTGAGGCCACCCCTGCGCGTCCCTGAGGAGGCC 541
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 18470 CCACCTTCTACTCCACATGAAATGGGGGAAACCTTGGCACTCCCTCAGGAGCAGC 18415
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
AC087162/c
LOCUS
DEFINITION Mus musculus clone RP23-232H18, *** SEQUENCING IN PROGRESS ***, 109
ACCESSION AC087162
VERSION AC087162.3 GI:13446280
KEYWORDS HTG; HTGS_PHASE1.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 207588)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Mus musculus, clone RP23-232H18
Unpublished
2 (bases 1 to 207588)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barra,N., Bastien,V., Boguslavsky,L., Boukhalter,B., Brown,A.,
Camaratte,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collimore,A., Cooke,P., Deakellian,K., Dewar,K., Diaz,J.S.,
Dodge,S., Fairo,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardyna,S., Glend,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., Larocque,K., Lamazares,R., Landers,T.,
Lenoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McKean,P., McKernan,K.,

```

TITLE
JOURNAL
COMMENT

MCPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Punthong, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roselli, M.,
Roy, A., Santos, R., Schauer, S., Schuppback, R., Seaman, S., Severy, P.,
Souarez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testa, S., Theodore, J.,
Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zalnoun, J.,
Zemke, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (09-DEC-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 25, 2001 this sequence version replaced gi:13123247.
All repeats were identified using RepeatMasker.
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L11590
Center clone name: 232_H18

NOTE: This is a 'working draft' sequence. It currently
consists of 109 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 545: contig of 545 bp in length
546 645: gap of 100 bp
646 822: contig of 177 bp in length
823 922: gap of 100 bp
923 1568: contig of 646 bp in length
1569 1668: gap of 100 bp
1669 1971: contig of 303 bp in length
1972 2071: gap of 100 bp
2072 3093: contig of 1022 bp in length
3094 3193: gap of 100 bp
3194 3721: contig of 528 bp in length
3722 3821: gap of 100 bp
3822 4856: contig of 1035 bp in length
4857 4956: gap of 100 bp
4957 5946: contig of 990 bp in length
5947 6046: gap of 100 bp
6047 6847: contig of 801 bp in length
6848 6947: gap of 100 bp
6948 7581: contig of 634 bp in length
7582 7681: gap of 100 bp
7682 8400: contig of 719 bp in length
8401 8500: gap of 100 bp
8501 9216: contig of 716 bp in length
9217 9316: gap of 100 bp
9317 9811: contig of 495 bp in length
9812 9911: gap of 100 bp
9912 10586: contig of 675 bp in length
10587 10686: gap of 100 bp
10687 11340: contig of 654 bp in length
11341 11440: gap of 100 bp
11441 12168: contig of 728 bp in length
12169 12268: gap of 100 bp
12269 12935: contig of 667 bp in length
12936 13035: gap of 100 bp
13036 13744: contig of 709 bp in length
13745 13844: gap of 100 bp
13845 14348: contig of 504 bp in length
14349 14448: gap of 100 bp

14449 15080: contig of 632 bp in length
15081 15180: gap of 100 bp
15181 15852: contig of 672 bp in length
15853 15952: gap of 100 bp
15953 16609: contig of 657 bp in length
16610 16709: gap of 100 bp
16710 17846: contig of 1137 bp in length
17847 17946: gap of 100 bp
17947 18571: contig of 625 bp in length
18572 18671: gap of 100 bp
18672 19518: contig of 847 bp in length
19519 19618: gap of 100 bp
19619 20250: contig of 632 bp in length
20251 20350: gap of 100 bp
20351 21033: contig of 683 bp in length
21034 21133: gap of 100 bp
21134 21558: contig of 425 bp in length
21559 21658: gap of 100 bp
21659 22346: contig of 688 bp in length
22347 22446: gap of 100 bp
22447 23720: contig of 1274 bp in length
23721 23820: gap of 100 bp
23821 24538: contig of 719 bp in length
24540 24639: gap of 100 bp
24640 25295: contig of 656 bp in length
25296 25395: gap of 100 bp
25396 26131: contig of 736 bp in length
26132 26231: gap of 100 bp
26232 26944: contig of 713 bp in length
26945 27044: gap of 100 bp
27045 27799: contig of 755 bp in length
27800 27899: gap of 100 bp
27900 28692: contig of 793 bp in length
28693 28792: gap of 100 bp
28793 29888: contig of 1096 bp in length
29889 29988: gap of 100 bp
29989 31437: contig of 1449 bp in length
31438 31537: gap of 100 bp
31538 32695: contig of 1158 bp in length
32696 32795: gap of 100 bp
32796 34017: contig of 1222 bp in length
34018 34117: gap of 100 bp
34118 35264: contig of 1147 bp in length
35265 35364: gap of 100 bp
35365 36141: contig of 777 bp in length
36142 36241: gap of 100 bp
36242 37401: contig of 1160 bp in length
37402 37501: gap of 100 bp
37502 38944: contig of 1443 bp in length
38945 39044: gap of 100 bp
39045 40236: contig of 1192 bp in length
40237 40336: gap of 100 bp
40337 41564: contig of 1228 bp in length
41565 41664: gap of 100 bp
41666 43110: contig of 1446 bp in length
43111 43210: gap of 100 bp
43211 44218: contig of 1008 bp in length
44219 44318: gap of 100 bp
44319 45382: contig of 1064 bp in length
45383 45482: gap of 100 bp
45483 46150: contig of 668 bp in length
46151 46250: gap of 100 bp
46251 47727: contig of 1477 bp in length
47728 47827: gap of 100 bp
47828 48840: contig of 1013 bp in length
48841 48940: gap of 100 bp
48941 50355: contig of 1415 bp in length
50356 50455: gap of 100 bp
50456 51763: contig of 1308 bp in length
51764 51863: gap of 100 bp
51864 53343: contig of 1480 bp in length
53344 53443: gap of 100 bp
53444 55038: contig of 1595 bp in length

```

* 55039 55138: gap of 100 bp
* 55139 56534: contig of 1396 bp in length
* 56535 56634: gap of 100 bp
* 56635 57911: contig of 1277 bp in length
* 57912 58011: gap of 100 bp
* 58012 58911: contig of 900 bp in length
* 58912 59011: gap of 100 bp
* 59012 60123: contig of 1112 bp in length
* 60124 60223: gap of 100 bp
* 60224 61343: contig of 1120 bp in length
* 61344 61443: gap of 100 bp
* 61444 63346: contig of 1903 bp in length
* 63347 63446: gap of 100 bp
* 63447 64248: contig of 802 bp in length
* 64249 64348: gap of 100 bp
* 64349 66453: contig of 2105 bp in length
* 66454 66553: gap of 100 bp
* 66554 67748: contig of 1195 bp in length
* 67749 67848: gap of 100 bp
* 67849 69558: contig of 1710 bp in length
* 69559 69658: gap of 100 bp
* 69659 70839: contig of 1181 bp in length
* 70840 70939: gap of 100 bp
* 70940 71912: contig of 973 bp in length
* 71913 72012: gap of 100 bp
* 72013 73882: contig of 1870 bp in length
* 73883 73982: gap of 100 bp
* 73983 75908: contig of 1926 bp in length
* 75909 76008: gap of 100 bp

```

Query Match
Best local similarity 59.5%; Score 40.8; DB 2; Length 207588;
Matches 69; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

```

QY 426 TACCTGAGATCTGGAGATGCTGAGTGTGGGGCCAGAGAACACACACTCACTG 485
Db 77663 TTCTGCGCTCAGGAGACTGCGGGGTGATGCTACTCACTCACTCACTCACTT 77604
QY 486 CCCACTTCATTCGTGTGACCTGCTGAGGCCACCTCCCTGCTGAGAGGCC 541
Db 77603 CCACCTTCTTACTCCACATGATGGCGGAGACTTGCCTCCCTGAGGAGC 77548

```

```

RESULT 15
AC018477/c 172090 bp DNA linear HTG 04-NOV-2000
LOCUS Homo sapiens chromosome 19 clone RP11-67A5, WORKING DRAFT SEQUENCE,
DEFINITION 14 unordered pieces.
AC018477
AC018477.12 GI:11079337
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 172090)
Muzny D.M., Adams C., Adio-Oduola B., Ali-Osman F.R., Allen C.,
Albrooks S.L., Amaratunga H.C., Are J.R., Banks T., Barabara J.,
Benton J., Blumage K., Blankenburg K., Bonnin D., Bouch J.,
Bowle S., Brileva M., Brown E., Brown M., Bryant N.P., Buhay C.,
Burche P., Burkett C., Burrell K.L., Byrd N.C., Cartron T.F.,
Carter M., Cavazos S.R., Chacko J., Chavez D., Chen G., Chen R.,
Chen Z., Chowdhury I., Christopoulos C., Cleveland C.D., Cox C.,
Coyle M.D., Dathorne S.R., David R., Davila M.L., Davis C.,
Davy-Carroll L., Dederich D.A., Delaney K.R., Delgado O.,
Denn A.L., Ding Y., Din H.H., Douthwaite K.J., Drepper H.,
Dugan-Rocha S., Durbin K.J., Earnhart C., Edgar D., Edwards C.C.,
Elhaj C., Escoto M., Falls T., Ferraguto D., Flaeg N., Ford J.,
Foster P., Frantz P., Gabisi A., Gao J., Garcia A., Garner T.,
Gara N., Gill R., Gorrell J.H., Guevara M., Gunaratne P., Hale S.,
Hamilton K., Harris C., Harris K., Hart M., Havlak P., Hayes A.,
Hernandez J., Hernandez O., Hodgson A., Hogues M., Holloway C.,
Hollins B., Homs F., Howard S., Huber J., Hulys S., Hume J.,

```

COMMENT

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korah, J., Kovar, C., Kratovic, J., Kuresh, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Lousseng, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., McElroy, M.P., Meador, M., Mel, G., Metzker, M., Miner, G., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenwo, S., Ogun, M., Okunolu, G., Oragunye, N., Oviedo, R., Parnes, E., Pu, L.L., Peery, U., Perez, L., Peters, L., Plickens, R., Plickens, R., Roite, M., Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shooshbari, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, J., Taylor, C., Taylor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, K., Vasquez, L., Vera, V., Villalón, D., Vanson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Wascyk, R., Wooden, S., Wallington, S., Williams, G., Williamson, A., Wleciyk, R., Woodson, D., Morley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 172090)
Morley, K.C.

Direct Submission
Submitted (12-DEC-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 3, 2000 this sequence version replaced gi:9929562.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HMML
Center clone name: RP11-67A5
----- Summary Statistics
Sequencing vector: M13, L08821
Chemistry: Dye-terminator Big Dye 100x of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 138579 bases at least Q40
Consensus quality: 157955 bases at least Q30
Consensus quality: 164360 bases at least Q20
Estimated insert size: 166402; sum-of-coverage estimation
Estimated insert size: 215352; agarose-fp estimation
Quality coverage: 2.9x in Q20 bases; agarose-fp estimation
Quality coverage: 3.8x in Q20 bases; sum-of-coverage estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 14 contigs. The true order of the pieces is
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	28263	contig of 28263 bp in length
28264	28363	gap of unknown length
28364	51869	contig of 23506 bp in length
51870	51969	gap of unknown length
51970	72904	contig of 20335 bp in length
72905	73005	gap of unknown length
73005	94296	contig of 21292 bp in length
94297	94396	gap of unknown length
94397	114901	contig of 20505 bp in length
114902	115001	gap of unknown length
115002	130564	contig of 15563 bp in length
130565	130664	gap of unknown length

* 130665 140193: contig of 9529 bp in length
* 140184 140293: gap of unknown length
* 140294 150299: contig of 10006 bp in length
* 150300 150399: gap of unknown length
* 150400 155530: contig of 5131 bp in length
* 155531 155630: gap of unknown length
* 155631 159162: contig of 3532 bp in length
* 159163 159262: gap of unknown length
* 159263 164868: contig of 5606 bp in length
* 164869 168208: gap of unknown length
* 168209 168308: contig of 3240 bp in length
* 168309 170810: gap of unknown length
* 170811 170910: contig of 2502 bp in length
* 170911 172090: gap of unknown length
Location/Qualifiers
Source 1. 172090
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="RP11-67A5"

BASE COUNT 41338 a 44179 c 43713 g 41606 t 1354 others
ORIGIN

Query Match
Best Local Similarity 5.3%; Score 40.4; DB 2; Length 172090;
Matches 155; Conservative 0; Mismatches 161; Indels 8; Gaps 1;
QY 154 GGGACAAGTCTACGACCCCTGCGACACTGTGTATGATGATGCGGTGCGCTGG 213
DB 36230 GGGAGATTCTCTGACCTCTGTCACACCCAGACACTCTCTGCGACCCACCATGG 36171
QY 214 CCAGACCCAGACGTGTGGAACCTGACCT-----TCAGAGTCTGCTTGGACAGTG 265
DB 36170 CCTCCCCGAGGTTCTGGAGCTTCACATCCTGAGCAGACAGTGGGCTCTGAGGGGATC 36111
QY 266 CTGCCCCGACCTTCATGATGGAAGCTGATAACAGAACTGCGACTGACCCGGACCTC 325
DB 36110 CAGCCCCGCGCCACCTCCAGCTCTCAGCTCCAGGCCAGAGCCCTTCCAGGGCTTC 36051
QY 326 GGATGACAGGCTTGTGCGAGTGCAGCTAATGGAACATCAGGGGAACGATGACTCTGG 385
DB 36050 CCATCCGCCATCATCTCTCTCCGTTGGGATCCAGGTTCAAGCAGACCTCAAGGCTGG 35991
QY 386 ATTCTCTTCTCTGGTGGGCTGGAGAAAGAGCTGTGTTACCTGAGATCTGGATGCT 445
DB 35990 AGCGTGTGCGCCGGAGATCGAGAGCAAGGGGCTATGTCAGGAGACTCCGACGA 35931
QY 446 GAGTGGCTGTTGGGGCCAGAGA 469
DB 35930 TCGTGAAGGTAAAGCGGGCAGACA 35907

Search completed: November 3, 2002, 00:56:02
Job time : 3133 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 22:59:17 ; Search time 212 Seconds

(without alignments)
6179.272 Million cell updates/sec

Title: US-09-092-297-5

Perfect score: 763
Sequence: 1 CTCACGTCGACACACCCAGA.....ATAATTATGACTTATATA 763

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_032802:*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	762.6	99.9	763	20	AAK01686
2	762.6	99.9	763	20	AAK01687
3	761	99.7	771	21	AAK77668
4	761	99.7	771	21	AAK37062
5	761	99.7	771	22	AAK521491
6	761	99.7	771	22	AAK85962
7	761	99.7	771	22	AAK54298
8	761	99.7	914	21	AAA08111
9	760.6	99.7	762	20	AAK24391

10	735	96.3	748	21	AAZ98158	Human signal pept
11	416.6	54.6	417	20	AAK01685	Urinary tract tiss
12	236	30.9	236	20	AAK01684	Urinary tract tiss
13	196	25.7	196	20	AAK01683	Urinary tract tiss
14	148.8	19.5	1338	23	AAK92056	DNA encoding novel
15	90.6	11.9	662	21	AAK58632	Human PRO982 prote
16	90.6	11.9	662	21	AAK51261	Human DNA encoding
17	90.6	11.9	662	21	AAK91476	Human PRO982 CDNA.
18	90.6	11.9	662	22	AAK91558	Human PRO982 CDNA.
19	88.8	11.6	378	22	AAK52127	Human AFP protein
20	78	10.2	146	23	AAK570346	DNA encoding novel
21	78	10.2	1286	23	AAK570301	DNA encoding novel
22	68	8.9	376	21	AAK45557	Human secreted exp
23	68	8.9	610	21	AAK37061	Human PRO1275 (UNQ
24	68	8.9	610	22	AAK521490	Human CDNA sequenc
25	68	8.9	610	22	AAK54297	DNA encoding prote
26	68	8.9	789	21	AAK87684	Human secreted pro
27	66.8	8.8	549	23	AAK592055	DNA encoding novel
28	64.8	8.5	445	22	AAK57123	Human foetal liver
29	64.8	8.5	445	22	AAK05787	Human bone marrow
30	64.8	8.5	445	22	AAK31411	Human bone marrow
31	63.8	8.5	445	22	AAK137303	Probe #5389 used t
32	63.8	8.4	245	22	ABA70342	Human foetal liver
33	63.8	8.4	245	22	AAK18578	Human brain expres
34	63.8	8.4	245	22	AAK44496	Human bone marrow
35	63.8	8.4	245	22	AAK50488	Probe #19174 used
36	58	7.6	1449	23	AAK74277	DNA encoding novel
37	58	7.6	1724	23	AAK77238	DNA encoding novel
38	58	7.6	1942	23	AAK77486	DNA encoding novel
39	58	7.6	2044	23	AAK69642	DNA encoding novel
40	42.8	5.6	378	16	AAK24159	Human gene signatu
41	39.8	5.2	10732	21	AAK10594	Gene encoding a su
42	39.4	5.2	467	22	ABA44931	Human breast cell
43	39.4	5.2	467	22	ABA55405	Human foetal liver
44	39.4	5.2	467	22	ABA25131	Probe #3597 for ge
45	39.4	5.2	467	22	AAK03648	Human brain expres

ALIGNMENTS

RESULT 1
AAK01686 standard; DNA; 763 BP.

21-APR-1999 (first entry)

Urinary tract tissue library BL172 gene full length clone 15548381H.

BL172: urinary tract; marker; cancer; recombinant; human; ss.

Homo sapiens.

MO9855656-A1.

10-DEC-1998.

05-JUN-1998; 98MO-US11693.

05-JUN-1997; 97US-0869579.

(ABBO) ABBOTT LAB.

Billing-medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J, Granadosen, Hodges SC, Klass MR, Kirochvil JD, Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;

WPI; 1999-045802/04.

New purified polynucleotide BL172 derivatives and encoded polypeptides - useful in the identification of markers, which are

PT Indicative of urinary tract diseases or conditions
 XX Claim 1; Page 89-90; 114pp; English.
 PS
 CC This represents a full-length sequence of a clone specific for urinary
 CC tract tissue library BL172. The BL172 polynucleotides (AA01683-87) are
 CC used to detect target BL172 polynucleotides and BL172 mRNA and the BL172
 CC polypeptides are used to detect BL172 antigens/antibodies in a test
 CC sample. The identification of certain markers in these methods are
 CC indicative of the presence of urinary tract disease, especially cancer.
 CC Additionally, the polypeptides are used to detect antibodies that bind
 CC specifically to a BL172 epitope derived from a BL172 polypeptide. Host
 CC cells containing a recombinant expression vector comprising the BL172
 CC polynucleotide sequences are used to produce the polypeptides containing
 CC BL172 epitopes. The methods aid in diagnosis, detection, staging,
 CC monitoring, prognostication, in vivo imaging, prevention and treatment
 CC of diseases or conditions associated with BL172, especially urinary
 CC tract cancer.
 CC
 XX
 SQ Sequence 763 BP; 162 A; 231 C; 195 G; 175 T; 0 other;

Query Match 99.9%; Score 762.6; DB 20; Length 763;
 Best Local Similarity 99.9%; Pred. No. 6,7e-215;
 Matches 762; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCACAGCAACCAAGAGCCATGCTCCCGAGGCTGATGCTGCTTTGCA 60
 DB 1 CTCACAGCAACCAAGAGCCATGCTCCCGAGGCTGATGCTGCTTTGCA 60
 QY 61 TTTCTGATGCTCAGGCTCTCTGCTCAGAGAGAGAGAGAGAGAGAGAGAG 120
 DB 61 TTTCTGATGCTCAGGCTCTCTGCTCAGAGAGAGAGAGAGAGAGAGAGAG 120
 QY 121 ACCTGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 DB 121 ACCTGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 QY 181 ACTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 DB 181 ACTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 QY 241 CCTCAGAGTCTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 DB 241 CCTCAGAGTCTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 QY 301 AGAAGTGGAGTCAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 DB 301 AGAAGTGGAGTCAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 QY 361 ACATGAGGAG 420
 DB 361 ACATGAGGAG 420
 QY 421 GGTGTACCTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 480
 DB 421 GGTGTACCTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 480
 QY 481 AACTGAGGAG 540
 DB 481 AACTGAGGAG 540
 QY 541 CCACAGGTCCTCTGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 600
 DB 541 CCACAGGTCCTCTGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 600
 QY 601 GGACTGAG 660
 DB 601 GGACTGAG 660
 QY 661 TGAGTGGAG 720
 DB 661 TGAGTGGAG 720

QY 721 CAGAGCAAGCAGAGATCATTAATTAATTAATTAATTAATTAATTAAT 763
 DB 721 CAGAGCAAGCAGAGATCATTAATTAATTAATTAATTAATTAATTAAT 763

RESULT 2
 AA01687
 ID AA01687 standard; DNA; 763 BP.
 AC AA01687;
 XX
 DT 21-APR-1999 (first entry)
 XX
 DE urinary tract tissue library BL172 gene consensus sequence.
 XX
 KM BL172; urinary tract; marker; cancer; recombinant; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 3..356
 FT /*tag= a
 FT /product= "BL172 polypeptide"
 FT /note= "the start codon is not indicated"

XX W09855656-A1.
 XX
 PD 10-DEC-1998.
 XX
 PF 05-JUN-1998; 98WO-US11693.
 XX
 PR 05-JUN-1997; 97US-0869579.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 PI Billings-medell PA, Cohen M, Colpitts TL, Friedman PN;
 PI Gordon J, Granadosen, Hodges SC, Klags MR, Kratochvil JD;
 PI Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;
 XX
 DR WPI: 1999-045802/04.
 XX
 P-PSDB: AAM92252.
 XX
 PT New purified polynucleotide BL172 derivatives and encoded
 PT polypeptides - useful in the identification of markers, which are
 PT indicative of urinary tract diseases or conditions
 PS
 PS Claim 1; Page 90; 114pp; English.
 XX
 CC This represents a consensus sequence of a clone specific for urinary
 CC tract tissue library BL172. The BL172 polynucleotides (AA01683-87) are
 CC used to detect target BL172 polynucleotides and BL172 mRNA and the BL172
 CC polypeptides are used to detect BL172 antigens/antibodies in a test
 CC sample. The identification of certain markers in these methods are
 CC indicative of the presence of urinary tract disease, especially cancer.
 CC Additionally, the polypeptides are used to detect antibodies that bind
 CC specifically to a BL172 epitope derived from a BL172 polypeptide. Host
 CC cells containing a recombinant expression vector comprising the BL172
 CC polynucleotide sequences are used to produce the polypeptides containing
 CC BL172 epitopes. The methods aid in diagnosis, detection, staging,
 CC monitoring, prognostication, in vivo imaging, prevention and treatment
 CC of diseases or conditions associated with BL172, especially urinary
 CC tract cancer.
 CC
 XX
 SQ Sequence 763 BP; 162 A; 231 C; 194 G; 175 T; 1 other;

Query Match 99.9%; Score 762.6; DB 20; Length 763;
 Best Local Similarity 100.0%; Pred. No. 6,7e-215;
 Matches 763; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCACAGCAACCAAGAGCCATGCTCCCGAGGCTGATGCTGCTTTGCA 60
 DB 1 CTCACAGCAACCAAGAGCCATGCTCCCGAGGCTGATGCTGCTTTGCA 60

PR	10-SEP-1998	9805-0099815
PR	10-SEP-1998	9805-0099816
PR	15-SEP-1998	9805-0100385
PR	15-SEP-1998	9805-0100386
PR	15-SEP-1998	9805-0100387
PR	15-SEP-1998	9805-0100388
PR	15-SEP-1998	9805-0100389
PR	16-SEP-1998	9805-0100547
PR	16-SEP-1998	9805-0100548
PR	16-SEP-1998	9805-0100661
PR	16-SEP-1998	9805-0100662
PR	16-SEP-1998	9805-0100664
PR	17-SEP-1998	9805-0100684
PR	17-SEP-1998	9805-0100685
PR	17-SEP-1998	9805-0100710
PR	17-SEP-1998	9805-0100711
PR	17-SEP-1998	9805-0100930
PR	17-SEP-1998	9805-0100939
PR	18-SEP-1998	9805-0100848
PR	18-SEP-1998	9805-0100849
PR	18-SEP-1998	9805-0100849
PR	18-SEP-1998	9805-0101014
PR	18-SEP-1998	9805-0101068
PR	18-SEP-1998	9805-0101071
PR	22-SEP-1998	9805-0101371
PR	22-SEP-1998	9805-0101372
PR	22-SEP-1998	9805-0101472
PR	22-SEP-1998	9805-0101473
PR	22-SEP-1998	9805-0101475
PR	23-SEP-1998	9805-0101476
PR	23-SEP-1998	9805-0101477
PR	23-SEP-1998	9805-0101478
PR	24-SEP-1998	9805-0101739
PR	24-SEP-1998	9805-0101741
PR	24-SEP-1998	9805-0101743
PR	24-SEP-1998	9805-0101743

PR	24-SEP-1998	9805-0101916
PR	24-SEP-1998	9805-0102207
PR	24-SEP-1998	9805-0102240
PR	29-SEP-1998	9805-0102320
PR	29-SEP-1998	9805-0102347
PR	29-SEP-1998	9805-0102350
PR	29-SEP-1998	9805-0102351
PR	30-SEP-1998	9805-0102487
PR	30-SEP-1998	9805-0102570
PR	30-SEP-1998	9805-0102571
PR	01-OCT-1998	9805-0102654
PR	01-OCT-1998	9805-0102687
PR	02-OCT-1998	9805-0102965
PR	06-OCT-1998	9805-0103258
PR	06-OCT-1998	9805-0103449
PR	07-OCT-1998	9805-0103314
PR	07-OCT-1998	9805-0103315
PR	07-OCT-1998	9805-0103335
PR	07-OCT-1998	9805-0103396
PR	07-OCT-1998	9805-0103401
PR	08-OCT-1998	9805-0103633
PR	08-OCT-1998	9805-0103679
PR	08-OCT-1998	9805-0103711
PR	14-OCT-1998	9805-0104257
PR	20-OCT-1998	9805-0104887
PR	20-OCT-1998	9805-0105000
PR	20-OCT-1998	9805-0105002
PR	21-OCT-1998	9805-0105104
PR	22-OCT-1998	9805-0105169
PR	22-OCT-1998	9805-0105263
PR	26-OCT-1998	9805-0105536
PR	26-OCT-1998	9805-0105594
PR	27-OCT-1998	9805-0105811
PR	27-OCT-1998	9805-0105882
PR	27-OCT-1998	9805-0106062
PR	28-OCT-1998	9805-0106023
PR	28-OCT-1998	9805-0106029

```

PR 28-OCT-1998; 98US-0106030.
PR 28-OCT-1998; 98US-0106032.
PR 28-OCT-1998; 98US-0106033.
PR 28-OCT-1998; 98US-0106178.
PR 28-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106384.
PR 29-OCT-1998; 98US-0108500.
PR 30-OCT-1998; 98US-0106464.
PR 03-NOV-1998; 98US-0106856.
PR 03-NOV-1998; 98US-0106902.
PR 03-NOV-1998; 98US-0106905.
PR 03-NOV-1998; 98US-0106919.
PR 03-NOV-1998; 98US-0106932.
PR 03-NOV-1998; 98US-0106934.
PR 10-NOV-1998; 98US-0107783.
PR 17-NOV-1998; 98US-0108775.
PR 17-NOV-1998; 98US-0108787.
PR 17-NOV-1998; 98US-0108788.
PR 17-NOV-1998; 98US-0108801.
PR 17-NOV-1998; 98US-0108802.
PR 17-NOV-1998; 98US-0108806.
PR 17-NOV-1998; 98US-0108807.
PR 17-NOV-1998; 98US-0108867.
PR 17-NOV-1998; 98US-0108925.
PR 18-NOV-1998; 98US-0108848.
PR 18-NOV-1998; 98US-0108849.
PR 18-NOV-1998; 98US-0108850.
PR 18-NOV-1998; 98US-0108851.
PR 18-NOV-1998; 98US-0108852.
PR 18-NOV-1998; 98US-0108858.
PR 18-NOV-1998; 98US-0108904.

XX (GETH ) GENENTECH INC.
XX
XX Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WT;
PI WPI: 2000-237871/20.
XX P-PSDB; AA199380.
XX
XX New mammalian DNA sequences encoding transmembrane, receptor or
PT secreted PRO polypeptides, useful for screening of potential peptide or
PR small molecule inhibitors of the relevant receptor/ligand interactions
XX
XX Claim 2: Fig 81; 773bp; English.
XX
XX AAA37022 to AAA37144 encode the new isolated human transmembrane,
CC receptor or secreted PRO polypeptides given in AA199340 to AA199462. The
CC transmembrane and receptor PRO proteins can be used for screening of
CC potential peptide or small molecule inhibitors of the relevant
CC receptor/ligand interactions. The polypeptides and nucleotide sequences
CC encoding them have various Industrial applications, including uses as
CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
CC PCR primers and hybridisation probes used in the isolation of the PRO
CC polypeptides from the present invention.
XX
XX .sequence 771 BP; 169 A; 231 C; 195 G; 176 T; 0 other;
SQ

Query Match 99.7%; Score 761; DB 21; Length 771;
Best Local Similarity 99.7%; Pred. No. 2e-214;
Matches 761; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CTCGACGTGACACCCAGAGCCATGGCTCCCGAGGCTGATCGTAGTCTTTGGCA 60
Db 1 CTCGACGTGACACCCAGAGCCATGGCTCCCGAGGCTGATCGTAGTCTTTGGCA 60
OY TTTTGTGATGTCGCGAGGTCCTGCTGTCACAGCGAGCCCGCAATGACTCTT 120
Db TTTTGTGATGTCGCGAGGTCCTGCTGTCACAGCGAGCCCGCAATGACTCTT 120
OY TTTTGTGATGTCGCGAGGTCCTGCTGTCACAGCGAGCCCGCAATGACTCTT 120
Db TTTTGTGATGTCGCGAGGTCCTGCTGTCACAGCGAGCCCGCAATGACTCTT 120
OY 121 ACGTGATGCTGTGCAGCAGACACAAGAGATGTGGGAGCAAGTTTACGACCCCGTCGAGC 180
Db 121 ACGTGATGCTGTGCAGCAGCAGACACAAGAGATGTGGGAGCAAGTTTACGACCCCGTCGAGC 180
OY 121 ACGTGATGCTGTGCAGCAGCAGACACAAGAGATGTGGGAGCAAGTTTACGACCCCGTCGAGC 180
Db 121 ACGTGATGCTGTGCAGCAGCAGACACAAGAGATGTGGGAGCAAGTTTACGACCCCGTCGAGC 180

```

OY	181	ACGTTGCTATGATGATGCCCGTCGTCCTTGGCCAGACCCAGACGTGTGGAAATGCA	240
Db	181	ACTGTTGCTATGATGATGCGCTGTCGTCCTTGGCCAGACCCAGACGTGTGGAAATGCA	240
OY	241	CCCTCAGAGTCTGCTTTGAGCAGATGCTGCGCCCTTGACCTTATGTTGAAGCTGATAACC	300
Db	241	CCCTCAGAGTCTGCTTTGAGCAGATGCTGCGCCCTTGACCTTATGTTGAAGCTGATAACC	300
OY	301	AGAACTGCGACTCAGCCCGGACCTCGATGACAGGCTTTGTGCGAGTGCAGCTAATGGA	360
Db	301	AGAACTGCGACTCAGCCCGGACCTCGATGACAGGCTTTGTGCGAGTGCAGCTAATGGA	360
OY	361	ACATCAGGGGGAACGATGATCCTGGAATTCCTCTTCCTGCTGGTGGGCTGGAAGAGGCT	420
Db	361	ACATCAGGGGGAACGATGATCCTGGAATTCCTCTTCCTGCTGGTGGGCTGGAAGAGGCT	420
OY	421	GGGTTCACCTGAGATCTGGAGTGTGAGTGGCTGTTGGGGGGCCAGAAACACACTC	480
Db	421	GGGTTCACCTGAGATCTGGAGTGTGAGTGGCTGTTGGGGGGCCAGAAACACACTC	480
OY	481	AACTGCCCATCTCAATCTGTGTGACACTGTCTGAGGCCCAACCTGCCGCTGAGGAGGC	540
Db	481	AACTGCCCATCTCAATCTGTGTGACACTGTCTGAGGCCCAACCTGCCGCTGAGGAGGC	540
OY	541	CCACAGTCTCCCTTTTAAATTCGAGACAGTATGATGCGTGTCTGATGGGGGCCAG	600
Db	541	CCACAGTCTCCCTTTTAAATTCGAGACAGTATGATGCGTGTCTGATGGGGGCCAG	600
OY	601	GGACTCTGAAACCTCTGATGACCCCTATGGCCAAATCAACCCGGCACACCCCAAGGC	660
Db	601	GGACTCTGAAACCTCTGATGACCCCTATGGCCAAATCAACCCGGCACACCCCAAGGC	660
OY	661	TGGCTGGGGAACCTTCAACCTTGTGTGAGATTTTCCATCATCTCAAGTTCCTTATC	720
Db	661	TGGCTGGGGAACCTTCAACCTTGTGTGAGATTTTCCATCATCTCAAGTTCCTTATC	720
OY	721	CAGGAGCAAGACAGGATCATATAATTTAGTACTTATA	763
Db	721	CAGGAGCAAGACAGGATCATATAATTTAGTACTTATA	763
RESULT 5			
ID	AA521491		
AC	AA521491 standard; cDNA; 771 bp.		
XX	AA521491;		
XX	24-OCT-2001 (first entry)		
XX	Human cDNA sequence encoding for PRO1274 polypeptide.		
KW	Human secretory and transmembrane; PRO; mammalian; cancer; lung;		
KW	breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;		
KW	catilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;		
XX	adipocyte; A-peptide; factor VIIA; gene therapy; ss.		
OS	Homo sapiens.		
XX			
XX	WO200140466-A2.		
XX	07-JUN-2001.		
XX			
XX	01-DEC-2000; 2000MO-US32678.		
XX			
PR	01-DEC-1999; 99MO-US28301.		
PR	01-DEC-1999; 99MO-US28634.		
PR	02-DEC-1999; 99MO-US28551.		
PR	02-DEC-1999; 99MO-US28564.		
PR	02-DEC-1999; 99MO-US28565.		
PR	09-DEC-1999; 99US-0170262.		
PR	16-DEC-1999; 99MO-US30095.		
PR	20-DEC-1999; 99MO-US30911.		

XX Key Location/Qualifiers
FH CDS 24..356
FT /tag- a
FT sig_peptide /product- "PRO1274 polypeptide"
FT 24..95
FT mat_peptide /tag- b
FT 96..353
FT /tag- c
PN WO200140465-A2.
XX 07-JUN-2001.
PD 10-NOV-2000; 2000WO-US30873.
XX 30-NOV-1999; 99WO-US28313.
PR 09-DEC-1999; 99US-0170262.
PR 23-DEC-1999; 99US-0172059.
PR 11-JAN-2000; 2000US-0175481.
PR 20-JAN-2000; 2000US-0177118.
PR 18-FEB-2000; 2000WO-US04342.
PR 03-MAR-2000; 2000US-0187202.
PR 30-MAY-2000; 2000WO-US14941.
PR 05-JUN-2000; 2000US-0209832.
PR 24-AUG-2000; 2000WO-US23328.
XX (GENE) GENENTECH INC.
XX Fong S, Goddard A, Godowski P, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Tunas D, Watanabe CK, Wood WI, Zhang Z;
XX WPI: 2001-381384/40.
DR P-PSDB: AAB47292.
XX Isolated PRO polypeptide useful for treat or diagnose an immune-related
XX disorder e.g. arthritis, asthma, allergy, diabetes or psoriasis -
XX
XX Claim 2; Fig 3; 124pp: English.
XX
XX The sequences given in AAC85961-69 encode PRO polypeptides. PRO1081,
CC PRO1274 and PRO10272 stimulate the proliferation of T-lymphocytes and
CC PRO1199, PRO1556, PRO4401 and PRO10266 inhibit the proliferation of
CC T-lymphocytes. PRO1754 and PRO9912 act to enhance the infiltration of
CC mononuclear cells, eosinophils or polymorphonuclear neutrophils (PMN)
CC into the tissue of a mammal. The PRO cDNA's and antibodies which
CC bind to them, are used to treat an immune-related disorder in a
CC mammal. Such disorders include systemic lupus erythematosus,
CC rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis, a
CC spondyloarthropathy, systemic sclerosis, an idiopathic inflammatory
CC myopathy, Sjogren's syndrome, systemic vasculitis, sarcoidosis,
CC autoimmune haemolytic anaemia, autoimmune thrombocytopaenia, thyroiditis,
CC diabetes mellitus, immune-mediated renal disease, a demyelinating
CC disease of the central or peripheral nervous system, idiopathic
CC demyelinating polyneuropathy, Guillain-Barre syndrome, a chronic
CC inflammatory demyelinating polyneuropathy, a hepatobiliary disease,
CC infectious or autoimmune chronic active hepatitis, primary biliary
CC cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, erythema multiforme, contact dermatitis, psoriasis, an
CC allergic disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease.
XX
XX Sequence 771 BP; 169 A; 231 C; 195 G; 176 T; 0 other:

QY 1 CTCCACTGCAACCAACCCAGAGCCATGGCTCCCGAGGCTGATGCTGTCTTGCA 60
Db 1 CTCACATGCAACCAACCCAGAGCCATGGCTCCCGAGGCTGATGCTGTCTTGCA 60
QY 61 TTTTCTGATCTCCAGGCTCTCTGCTCACAGGAGCCCAAGTGGCCCAATGACTCTT 120
Db 61 TTTTCTGATCTCCAGGCTCTCTGCTCACAGGAGCCCAAGTGGCCCAATGACTCTT 120
QY 121 ACCGATGCTGTGCGACGACACAGAGATGTGGGACAACTTTACGACCCCTGACG 180
Db 121 ACCGATGCTGTGCGACGACACAGAGATGTGGGACAACTTTACGACCCCTGACG 180
QY 181 ACTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
Db 181 ACTGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
QY 241 CCTTCAGAGTCTGCTTGTGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 241 CCTTCAGAGTCTGCTTGTGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 301 AGAATCTGCACTGACCCCGACCTCGATGACAGGCTTTTCCAGTGTGCTAATGA 360
Db 301 AGAATCTGCACTGACCCCGACCTCGATGACAGGCTTTTCCAGTGTGCTAATGA 360
QY 361 ACATCAGGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 361 ACATCAGGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 421 GGTGTACCTGAGATCTGGGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 421 GGTGTACCTGAGATCTGGGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 AACTGCCCATTTATTTGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 481 AACTGCCCATTTATTTGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 541 CCAGAGTCTCTTCTAGATTTCTGACAGCAGATGATGATGATGATGATGATGATG 600
Db 541 CCAGAGTCTCTTCTAGATTTCTGACAGCAGATGATGATGATGATGATGATGATG 600
QY 601 GGAATCTGAACCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db 601 GGAATCTGAACCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 TGGCTGGGGAACCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 661 TGGCTGGGGAACCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 721 CAGGACCAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 763
Db 721 CAGGACCAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 763
RESULT 7
AAF54298
ID AAF54298 standard; DNA; 771 BP.
XX
AC AAF54298;
XX
DT 02-APR-2001 (first entry)
XX
DE DNA encoding protein of the invention #41.
XX
KW Secreted; transmembrane; gene therapy; ss.
XX
OS unidentified.
XX
PN WO200078961-A1.
XX
PD 28-DEC-2000.
XX
PF 18-FEB-2000; 2000WO-US04342.
XX

PR	23-JUN-1999;	99US-0141037.
PR	20-JUL-1999;	99US-0144758.
PR	26-JUL-1999;	99US-0145698.
PR	01-SEP-1999;	99MO-US20111.
PR	29-OCT-1999;	99US-0162506.
PR	30-NOV-1999;	99MO-US28313.
PR	02-DEC-1999;	99MO-US28551.
PR	16-DEC-1999;	99MO-US30095.
PR	05-JAN-2000;	2000MO-US00219.
PR	06-JAN-2000;	2000MO-US00376.
XX		
PA	(GETH) GENENTECH INC.	
XX		
P1	Baker KP, Bolstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S,	
P1	Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;	
P1	Pan J, Peoni NF, Roy MA, Smith V, Stewart TA, Tumas D;	
P1	Metanabe CK, Williams PM, Wood WI;	
XX		
XX	WPI: 2001-071395/08.	
XX		
PT	Secreted and transmembrane proteins and nucleic acids designated PRO,	
PT	useful as hybridization probes, in chromosome and gene mapping and gene	
PT	therapy -	
XX		
PS	Claim 2; Fig 8I; 787pp; English.	
XX		
CC	The present invention relates to secreted and transmembrane proteins.	
CC	These proteins and the DNA encoding them may be used as hybridization	
CC	probes, in chromosome and gene mapping and in the generation of	
CC	anti-sense RNA and DNA. They may also be used used to generate either	
CC	transgenic animals or knockout animals which are in turn useful for	
CC	development and screening of therapeutically useful reagents.	
CC	The nucleic acids may also be used in gene therapy.	
XX		
SQ	Sequence 771 BP; 169 A; 231 C; 195 G; 176 T; 0 other;	
Query Match	99.7%; Score 761; DB 22; Length 771;	
Best Local Similarity	99.7%; Fred. No. 2e-214;	
Matches 761; Conservative 1; Mismatches 1; Indels 0; Gaps 0;		
QY	1 CTCACATGCAACACCAGCAGCGTGGCCCCGAGGCGTCATCGTAGTGCTTTGGCA 60	
DB	1 CTCACATGCAACACCAGCAGCGTGGCCCCGAGGCGTCATCGTAGTGCTTTGGCA 60	
QY	61 TTTTGTGCATCTCCAGGCTCCTGTGCTCACAGCAGGCCCCAGTGGCCCCCATGACTCTT 120	
DB	61 TTTTGTGCATCTCCAGGCTCCTGTGCTCACAGCAGGCCCCAGTGGCCCCCATGACTCTT 120	
QY	121 ACCGTATGCTGTGCCAGCCACAAGAAGATGTGGGACAAAGTTTACACACCCCCTGCAC 180	
DB	121 ACCGTATGCTGTGCCAGCCACAAGAAGATGTGGGACAAAGTTTACACACCCCCTGCAC 180	
QY	181 ACTGTGTATGATATCCGCGTCGCGCCCTTGGAGGACCCAGACCGATGTGAAGAACTGCA 240	
DB	181 ACTGTGTATGATATCCGCGTCGCGCCCTTGGAGGACCCAGACCGATGTGAAGAACTGCA 240	
QY	241 CCTTCAGAGTGTCTTTGAGCAGTGTGCTGCCCTTGACCTTCATGTGTGAAGCTGATAAAC 300	
DB	241 CCTTCAGAGTGTCTTTGAGCAGTGTGCTGCCCTTGACCTTCATGTGTGAAGCTGATAAAC 300	
QY	301 AGAAGTGGACTCACACCCCGGACCTGGGATGACAGGCTTTGTGCAAGTGCACCTAATGGA 360	
DB	301 AGAAGTGGACTCACACCCCGGACCTGGGATGACAGGCTTTGTGCAAGTGCACCTAATGGA 360	
QY	361 ACATCAGGGAAGCATGTAAGTCTGTGATTCCTCTTCTGGTGGGCTGGAGAAAAGAGCT 420	
DB	361 ACATCAGGGAAGCATGTAAGTCTGTGATTCCTCTTCTGGTGGGCTGGAGAAAAGAGCT 420	
QY	421 GGGTGTACCTGAGATCTGGAGTGTGTAAGTGTGTTTGGGGCCAGAGAAACACACACTC 480	
DB	421 GGGTGTACCTGAGATCTGGAGTGTGTAAGTGTGTTTGGGGCCAGAGAAACACACACTC 480	
QY	481 AACGCGCACATTCTGTGTAAGCTGTGTGAAGCCACCCCTGCGCTGCTGAGAGAGCC 540	

Db	481	AACTGCCACTTCATTTCTGTGACCTGTCTGAGGCCACCCTCAGCTGCGCTGAGAGGC	540
Qy	541	CCACAGGTCCTCTTAGAATTCCTGGACAGCATGAGATCGTGTGATGGGGCCCG	600
Db	541	CCACAGGTCCTCTTAGAATTCCTGGACAGCATGAGATCGTGTGATGGGGCCCG	600
Qy	601	GGACTCTGAACTCTCTGATGACCCSTATGCGCAACATCACCCGGCACCCCAAGGC	660
Db	601	GGACTCTGAACTCTCTGATGACCCSTATGCGCAACATCACCCGGCACCCCAAGGC	660
Qy	661	TGGCTGGGGAACCTTCACCTCTGTGATGATTTCCATCATCTCAAGTCTCTTATC	720
Db	661	TGGCTGGGGAACCTTCACCTCTGTGATGATTTCCATCATCTCAAGTCTCTTATC	720
Qy	721	CAGGAGCAAGACACAGATCATATAATTTATGACTTTATA	763
Db	721	CAGGAGCAAGACACAGATCATATAATTTATGACTTTATA	763
RESULT 8			
XX	AAA08111		
XX	AAA08111	standard; DNA; 914 BP.	
XX	AAA08111;		
XX	22-JUN-2000	(first entry)	
XX	Human bladder specific gene 22638	clone 819141H SEQ ID NO:1.	
XX	Human; bladder specific gene; BSG; bladder cancer; diagnosis;		
XX	detection; metastasis; cytostatic; ds.		
XX	Homo sapiens.		
XX	WO200012761-A1.		
XX	09-MAR-2000.		
XX	01-SEP-1999;	99WO-US20003.	
XX	02-SEP-1998;	98US-0098825.	
XX	(DIAD-)	DIADEXUS LLC.	
XX	Salceda S, Sun Y, Reclon H, Cafierkey R;		
XX	WPI; 2000-256658/22.		
XX	Detecting, diagnosing metastasis, staging, monitoring, imaging and		
XX	treating bladder cancer. Involves measuring bladder specific gene (BSG)		
XX	levels in cells or body fluids		
XX	Claim 6; Page 25-26; 31pp; English.		
XX	A method has been developed for detecting, diagnosing metastasis, staging and monitoring the onset of metastasis or a change in stage of bladder cancer, by measuring bladder specific gene (BSG) levels in cells, tissues or body fluids. An increase in BSG levels compared to normal human control is associated with the presence or progression of bladder cancer, a decrease is associated with cancer regression. The method comprising measuring BSG is useful for detecting, diagnosing metastasis, staging and monitoring the onset of metastasis or a change in stage of bladder cancer. An antibody against BSG (I) labeled with paramagnetic ions or an isotope is useful for imaging bladder cancer and (II), especially conjugated to a cytotoxic agent is useful for treating bladder cancer. The method is efficient in discriminating between bladder cancer which has metastasised and the bladder cancer which has not metastasised. The present sequence represents a BSG sequence used in the exemplification of the present invention.		
XX	Sequence 914 BP; 193 A; 265 C; 237 G; 218 T; 1 other;		

Query Match 99.7%; Score 761; DB 21; Length 914;
 Best Local Similarity 99.7%; Pred. No. 2.2e-214;
 Matches 761; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 CTCACACTGCAACCAACCCAGACCATGGCTCCCGAGAGCTGATGATGCTTTGGCA 60
DB 33 CTCACACTGCAACCAACCCAGACCATGGCTCCCGAGAGCTGATGATGCTTTGGCA 92
OY 61 TTTTTCGATCTCCAGAGCTCTCTGCTCACACGAGAGCCAGTGGCCCATGACTCTT 120
DB 93 TTTTTCGATCTCCAGAGCTCTCTGCTCACACGAGAGCCAGTGGCCCATGACTCTT 152
OY 121 ACTGTGATGGTGGCCAGCCACACAGAGATGGGGGACAGTGTACAGACCCCTGACG 180
DB 153 ACTGTGATGGTGGCCAGCCACACAGAGATGGGGGACAGTGTACAGACCCCTGACG 212
OY 181 ACTGTGATGGTGGCCAGCCAGCTGCTGCTGGCCAGAGAGCCAGAGCTGTGAAACGCA 240
DB 213 ACTGTGATGGTGGCCAGCCAGCTGCTGCTGGCCAGAGAGCCAGAGCTGTGAAACGCA 272
OY 241 CCTTCAGAGTCTGCTTTGAGACAGTGTGCTGCCCTGAGACCTTCATGTGTAAGCTGTAAC 300
DB 273 CCTTCAGAGTCTGCTTTGAGACAGTGTGCTGCCCTGAGACCTTCATGTGTAAGCTGTAAC 332
OY 301 AGAAGTCGACATCAGCCCGAGCTCGATGACAGAGCTTTTCCGAGTGTACGTAATGGA 360
DB 333 AGAAGTCGACATCAGCCCGAGCTCGATGACAGAGCTTTTCCGAGTGTACGTAATGGA 392
OY 361 ACATCAGGGGAGAGATGACTCTGATGATCTGCTGCTGGTGGGGCCGGAAGAGGCT 420
DB 393 ACATCAGGGGAGAGATGACTCTGATGATCTGCTGCTGGTGGGGCCGGAAGAGGCT 452
OY 421 GGTGTACCTGAGATCTGGGATGCTAGTGGCTGTTGGGGGCGACAGAAACACACTC 480
DB 453 GGTGTACCTGAGATCTGGGATGCTAGTGGCTGTTGGGGGCGACAGAAACACACTC 512
OY 481 AACTGCGCACTTCATCTGATGACCTGTCGAGAGCCCACTGCGCTGCGCTGAGAGAGG 540
DB 513 AACTGCGCACTTCATCTGATGACCTGTCGAGAGCCCACTGCGCTGCGCTGAGAGAGG 572
OY 541 CCACAGAGTCCCTTTAGATTTCTGACACAGATGAGATGCTGCTGATGGGGGCGCA 600
DB 573 CCACAGAGTCCCTTTAGATTTCTGACACAGATGAGATGCTGCTGATGGGGGCGCA 632
OY 601 GCACACTGTAACCTCTCTGATGACCCGTATGGCCACATCAACCCGCGACACCCCAAGC 660
DB 633 GCACACTGTAACCTCTCTGATGACCCGTATGGCCACATCAACCCGCGACACCCCAAGC 692
OY 661 TGGCGGGGGAACCTTCACCCCTTGTGTGAGATTTTTCATCATCTCAAGTTCTTCTATC 720
DB 693 TGGCGGGGGAACCTTCACCCCTTGTGTGAGATTTTTCATCATCTCAAGTTCTTCTATC 752
OY 721 CAGGAGCAAAAGCAGAGATCAATTAATTTTATGACTTATTA 763
DB 753 CAGGAGCAAAAGCAGAGATCAATTAATTTTATGACTTATTA 795

```

RESULT 9
 AA224391
 ID AA224391 standard; cDNA; 762 BP.
 AC AA224391;
 XX
 XX 14-FEB-2000 (first entry)
 DE Human bladder tumour cDNA library derived EST 3.
 XX
 XX Expressed sequence tag; human; bladder; tumour; cancer; cytostatic;
 KW Treatment; gene therapy; EST; ss.
 XX Homo sapiens.
 XX
 PN DE19818619-A1.

XX 28-OCT-1999.
 PD 21-APR-1998; 98DE-1018619.
 XX 21-APR-1998; 98DE-1018619.
 PF 21-APR-1998; 98DE-1018619.
 XX
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
 PA Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pillarsky C, Dahl E;
 PI WPI; 1999-612028/53.
 DR
 XX New nucleic acid sequences expressed in bladder tumor tissue, and
 PT derived polypeptides, for treatment of bladder tumor and identification
 of therapeutic agents
 PS
 XX Claim 3; Page 62; 132pp; German.

This invention describes novel polypeptide fragments (I) and the
 CC polynucleotides (II) that encode them that are highly expressed in a
 CC human bladder tumour and which have cytostatic activity. (II) are used
 CC for recombinant expression of (I) and to isolate complete genes. (I) are
 CC used to identify agents suitable for treatment of bladder cancer, to
 CC directly treat this form of cancer (including expression from gene
 CC therapy vectors) or are used in a preparation for cancer treatment. (I)
 CC is also used for the generation of specific antibodies. (II) are
 CC identified by assembling ESTs (expressed sequence tags) from a
 CC particular tissue type before comparison of expression patterns. This
 CC allows a significantly longer fragment of the gene to be revealed, and
 CC therefore reduces the number of failures associated with the fact that
 CC ESTs from different libraries may represent different parts of the same
 CC unknown gene, distorting the estimated frequency of occurrence in a
 CC particular tissue. AA24360-243309 represent expressed sequence tag (EST)
 CC fragments isolated from a human bladder tumour cDNA library which encode
 CC the proteins represented in AA26143-Y6198.
 XX
 SQ Sequence 762 BP; 162 A; 232 C; 194 G; 174 T; 0 other;

Query Match 99.7%; Score 760.6; DB 20; Length 762;
 Best Local Similarity 99.9%; Pred. No. 2.6e-214;
 Matches 760; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 CTCACACTGCAACCAACCCAGACCATGGCTCCCGAGAGCTGATGATGCTTTGGCA 60
DB 1 CTCACACTGCAACCAACCCAGACCATGGCTCCCGAGAGCTGATGATGCTTTGGCA 60
OY 61 TTTTTCGATCTCCAGAGCTCTCTGCTCACACGAGAGCCAGTGGCCCATGACTCTT 120
DB 61 TTTTTCGATCTCCAGAGCTCTCTGCTCACACGAGAGCCAGTGGCCCATGACTCTT 120
OY 121 ACTGTATGCTGTGCGCAGCCACACAGAGATGGGGGACAAAGTTCTTACAGACCCCTGACG 180
DB 121 ACTGTATGCTGTGCGCAGCCACACAGAGATGGGGGACAAAGTTCTTACAGACCCCTGACG 180
OY 181 ACTGTGATGGTGGCCAGCCAGCTGCTGCTGGCCAGAGAGCCAGAGCTGTGAAACGCA 240
DB 181 ACTGTGATGGTGGCCAGCCAGCTGCTGCTGGCCAGAGAGCCAGAGCTGTGAAACGCA 240
OY 241 CCTTCAGAGTCTGCTTTGAGACAGTGTGCTGCCCTGAGACCTTCATGTGTAAGCTGTAAC 300
DB 241 CCTTCAGAGTCTGCTTTGAGACAGTGTGCTGCCCTGAGACCTTCATGTGTAAGCTGTAAC 300
OY 301 AGAAGTCGACATCAGCCCGAGCTCGATGACAGAGCTTTTCCGAGTGTACGTAATGGA 360
DB 301 AGAAGTCGACATCAGCCCGAGCTCGATGACAGAGCTTTTCCGAGTGTACGTAATGGA 360
OY 361 ACATCAGGGGAGAGATGACTCTGATGATCTGCTGCTGGTGGGGCCGGAAGAGGCT 420
DB 361 ACATCAGGGGAGAGATGACTCTGATGATCTGCTGCTGGTGGGGCCGGAAGAGGCT 420
OY 421 GGTGTACCTGAGATCTGGGATGCTAGTGGCTGTTGGGGGCGACAGAAACACACTC 480

```

Db 421 GGTGTACCTGAGATCTGGATGCTGAGTGGCTGTTGGGGCCAGAGAAACACACACTC 480
 QY 481 AACTGCCACATTCATTTCTGTGACCTGTCTGAGGCCACCTGCGCTCCCTGAGAGGC 540
 Db 481 AACTGCCACATTCATTTCTGTGACCTGTCTGAGGCCACCTGCGCTCCCTGAGAGGC 540
 QY 541 CCACAGGTCCTTCTAGATTTCTGACAGCATGAGATGCTGTCTATGAGGGCCAG 600
 Db 541 CCACAGGTCCTTCTAGATTTCTGACAGCATGAGATGCTGTCTATGAGGGCCAG 600
 QY 601 GGACTCTGACCTCTGTGATGACCTGTAGGCGCAATCAACCCGGACACCCCAAGC 660
 Db 601 GGACTCTGACCTCTGTGATGACCTGTAGGCGCAATCAACCCGGACACCCCAAGC 660
 QY 661 TGGCTGGGGAACCTTCACCTTCTGTGAGATTTTCATCATCTCAAGTCTCTATC 720
 Db 661 TGGCTGGGGAACCTTCACCTTCTGTGAGATTTTCATCATCTCAAGTCTCTATC 720
 QY 721 CAGGAGCAAGCAGCATCAATAAATTATGACTTTA 761
 Db 721 CAGGAGCAAGCAGCATCAATAAATTATGACTTTA 761

RESULT 10

AA298158 standard; cDNA; 748 BP.

AA298158;

11-MAY-2000 (first entry)

Human signal peptide containing protein HSPB-50 cDNA SEQ ID NO:184.

Human; signal peptide-containing protein; HSP; diagnosis; cancer;
 inflammation; cardiovascular disease; anticancer; anti-inflammation;
 antimicrobial; neuroprotective; cardiovascular; hepatotropic;
 antiasthmatic; gene therapy; cell proliferation; neurological disorder;
 reproductive disorder; developmental disorder; arteriosclerosis;
 cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;
 asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;
 Parkinson's disease; Huntington's disease; ovulatory defect;
 muscular dystrophy; ss.

Homo sapiens.

WO200000610-A2.

06-JAN-2000.

25-JUN-1999; 99WO-US14484.

26-JUN-1998; 98US-0090762.

31-JUL-1998; 98US-0094983.

01-OCT-1998; 98US-0102686.

11-DEC-1998; 98US-0112129.

(INCY-) INCYTE PHARM INC.

Lai P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR;

Akerblom IE, Au-Young J, Yee H, Patterson C, Reddy R, Hillman JL;

Bandman O;

WPI: 2000-160673/14.

P-PSDB: AAY87273.

New human signal peptide-containing proteins useful in treatment, prevention and diagnosis of e.g. cancer, inflammation and cardiovascular disease

Claim 9; Page 281; 327pp; English.

AA298109 to AA298242 encode AAY87224 to AAY87357 which represent the human signal peptide-containing proteins HSPB-1 to HSPB-134. HSPBs have

CC anticancer, anti-inflammatory, antimicrobial, nootropic, hepatotropic,
 CC neuroprotective, cardiovascular and antiasthmatic activities, and can
 CC be used in gene therapy. HSPBs can be used to treat or prevent disorders
 CC associated with decreased activity or function of HSP. Antagonists of
 CC HSP are used to treat or prevent disorders associated with increased
 CC activity or function of HSP. Such diseases include cell proliferation
 CC (including cancer), inflammation, cardiovascular, neurological,
 CC reproductive or developmental disorders, (e.g. arteriosclerosis,
 CC cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,
 CC asthma, Crohn's disease, microbial or other infections, congestive or
 CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's
 CC diseases, schizophrenia, ovulatory defects, muscular dystrophy. HSP
 CC nucleic acids can be used for the recombinant production of HSP, for
 CC detecting HSP in standard hybridisation and amplification assays (for
 CC diagnosis and monitoring), in gene therapy, as antisense,
 CC triplex-forming or ribozyme therapeutics, for detecting related sequences
 CC or genetic variations, and for chromosomal mapping. HSP are also used to
 CC raise specific antibodies (Ab) and to screen for agonists and
 CC antagonists (potential therapeutic agents). Ab are used to diagnose, or
 CC monitor, HSP-related diseases (in usual immunoassays), as therapeutic
 CC antagonists, in competitive drug screens, and for purification of HSP
 CC from natural sources.

Sequence 748 BP; 159 A; 227 C; 193 G; 169 T; 0 other;

Query Match 96.3%; Score 735; DB 21; Length 748;

Best Local Similarity 99.6%; Pred. No. 9.3e-207; Matches 746; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 6 CTGCACACCCAGAGCATGCTCCCGAGGCTGATGCTGCTTTGGCATTTTC 65
 Db 1 CTGCACACCCAGAGCATGCTCCCGAGGCTGATGCTGCTTTGGCATTTTC 60
 QY 66 TGCATCTCAGAGCTCTCTGTCACAGAGCCCGAGGGCCCGCATGCTTAACTG 125
 Db 61 TGCATCTCAGAGCTCTCTGTCACAGAGCCCGAGGGCCCGCATGCTTAACTG 120
 QY 126 ATGCTGTGCCAGCAGCAGAGAGATGTGGGAGCAAGTCTAGACCCCTGAGCATGT 185
 Db 121 ATGCTGTGCCAGCAGCAGAGAGATGTGGGAGCAAGTCTAGACCCCTGAGCATGT 180
 QY 186 TGTATGATGATGCCGTGCTGCTGCTGGCCAGAGCCAGAGGTGTGAACCTGACCTTC 245
 Db 181 TGTATGATGATGCCGTGCTGCTGCTGGCCAGAGCCAGAGGTGTGAACCTGACCTTC 240
 QY 246 AGAGTCTGCTTGAAGAGTGTGCTGCCCTGGAGCTTCAATGTAAGCTATAACAGAGAC 305
 Db 241 AGAGTCTGCTTGAAGAGTGTGCTGCCCTGGAGCTTCAATGTAAGCTATAACAGAGAC 300
 QY 306 TGCAGCTAGCCCGAGCTCGAGTACAGAGCTTTGTGACAGTGTGCTAATGAGCAATC 365
 Db 301 TGCAGCTAGCCCGAGCTCGAGTACAGAGCTTTGTGACAGTGTGCTAATGAGCAATC 360
 QY 366 AGGGAGAGATGACTCTGATTTCTCTTCTGCTGGGTGGGCTGAGAAAGAGCTGTGT 425
 Db 361 AGGGAGAGATGACTCTGATTTCTCTTCTGCTGGGTGGGCTGAGAAAGAGCTGTGT 420
 QY 426 TACCTGAGATGTGGAGTGTAGAGGCTGTGGGGGCGAGAGAAACACACATCAATG 485
 Db 421 TACCTGAGATGTGGAGTGTAGAGGCTGTGGGGGCGAGAGAAACACACATCAATG 480
 QY 486 CCACTTCATCTGTGACCTGTCTGAGGCCACCTGCGCTGCTGAGAGGCCACCA 545
 Db 481 CCACTTCATCTGTGACCTGTCTGAGGCCACCTGCGCTGCTGAGAGGCCACCA 540
 QY 546 GGTCCCTTCTAGAAATTTCTGACAGCATGAGATGCTGTGATGGGGCCAGGACT 605
 Db 541 GGTCCCTTCTAGAAATTTCTGACAGCATGAGATGCTGTGATGGGGCCAGGACT 600
 QY 606 CTGAACCTCTGATGACCCGTATGGCCCAATCAACCCGGACACCCCAAGGCTGGT 665
 Db 601 CTGAACCTCTGATGACCCGTATGGCCCAATCAACCCGGACACCCCAAGGCTGGT 660

QY 666 GGGAGACCTTCACCTTCGTGAGATTTCCATCATCTCAAGTCTCTTCAATCCAGGA 725
 DB 661 -GGGAGACCTTCACCTTCGTGAGATTTCCATCATCTCAAGTCTCTTCAATCCAGGA 719
 QY 726 GCAGACACAGATCATTAATTAATTTATG 754
 DB 720 GCAGACACAGATCATTAATTAATTTATG 748

RESULT 11

AA01685/c
 ID AA01685 standard; DNA; 417 BP.

AC AA01685;
 XX 21-APR-1999 (first entry)

DE Urinary tract tissue library BL172 gene specific cDNA clone g2178680.
 XX BL172; urinary tract; marker; cancer; recombinant; human; ss.

OS Homo sapiens.

PN W09855656-A1.

PD 10-DEC-1998.

PE 05-JUN-1998; 98WO-US11693.

PR 05-JUN-1997; 97US-0869579.

PA (ABBO) ABBOTT LAB.

PI Billing-medel PA, Cohen M, Colpitts TL, Friedman PN;
 PI Gordon J, Granadosen, Hodges SC, Klass MR, Kratochvil JD;
 PI Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;

DR WPI: 1999-045802/04.

PT New purified polynucleotide BL172 derivatives and encoded
 PT polypeptides - useful in the identification of markers, which are
 PT indicative of urinary tract diseases or conditions

PS Claim 1; Page 89; 114pp; English.

CC Sequences AA01683-85 represent overlapping clones specific for urinary
 CC tract tissue library BL172. The BL172 polynucleotides are used to detect
 CC target BL172 polynucleotides and BL172 mRNA and the BL172 polypeptides
 CC are used to detect BL172 antigens/antibodies in a test sample. The
 CC identification of certain markers in these methods are indicative of the
 CC presence of urinary tract disease, especially cancer. Additionally, the
 CC BL172 epitope derived from a BL172 polypeptide. Host cells containing a
 CC recombinant expression vector comprising the BL172 polynucleotide
 CC sequences are used to produce the polypeptides containing BL172
 CC epitopes. The methods aid in diagnosis, detection, staging, monitoring,
 CC prognostication, in vivo imaging, prevention and treatment of diseases
 CC or conditions associated with BL172, especially urinary tract cancer.

XX Sequence 417 BP; 99 A; 108 C; 117 G; 93 T; 0 other;

QY Query Match 54.6%; Score 416.6; DB 20; Length 417;
 DB Best Local Similarity 99.8%; Pred. No. 6.6e-113;
 DB Matches 416; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 346 GTGTCACTATGAGATCAGGAGAGATGACTCTGTGATTCCTCTCTGCTGGTGGC 405

DB 417 GTGTCACTATGAGATCAGGAGAGATGACTCTGTGATTCCTCTCTGCTGGTGGC 358

QY 406 CTGAGAGAGAGCTGTGTTTACCTGAGATCTGGAGTCTGAGTGTGTTGGGGCCA 465
 DB 357 CTGAGAGAGAGCTGTGTTTACCTGAGATCTGGAGTCTGAGTGTGTTGGGGCCA 298

QY 466 GAGAAACACACTCAACTGCCACTTCATTCGTGACCTGTGAGGCCACCTGGCCG 525
 DB 297 GAGAAACACACTCAACTGCCACTTCATTCGTGACCTGTGAGGCCACCTGGCCG 238
 QY 526 CTGCCCTGAGAGAGCCACAGGTCCTTGTGAGATTTGAGACAGATGATGCTGTG 585
 DB 237 CTGCCCTGAGAGAGCCACAGGTCCTTGTGAGATTTGAGACAGATGATGCTGTG 178
 QY 586 CTGATGGGGGCCAGGAGACTGTAACCTTCGTGATACCCCTATGGCCACATCAACCCG 645
 DB 177 CTGATGGGGGCCAGGAGACTGTAACCTTCGTGATACCCCTATGGCCACATCAACCCG 118
 QY 646 GCACACACCCCAAGGCTGCTGGGAGAACCTTCACCTTCGTGAGATTTTCATCATCTC 705
 DB 117 GCACACACCCCAAGGCTGCTGGGAGAACCTTCACCTTCGTGAGATTTTCATCATCTC 58
 QY 706 AAGTCTCTTCTATCCAGAGACAGACAGATCATTAATTAATTTATGACTTTAT 762
 DB 57 AAGTCTCTTCTATCCAGAGACAGACAGATCATTAATTAATTTATGACTTTAT 1

RESULT 12

AA01684
 ID AA01684 standard; DNA; 236 BP.

AC AA01684;

DE 21-APR-1999 (first entry)

DE Urinary tract tissue library BL172 gene specific cDNA clone 2624118.

XX BL172; urinary tract; marker; cancer; recombinant; human; ss.

OS Homo sapiens.

PN W09855656-A1.

PD 10-DEC-1998.

PE 05-JUN-1998; 98WO-US11693.

PR 05-JUN-1997; 97US-0869579.

PA (ABBO) ABBOTT LAB.

PI Billing-medel PA, Cohen M, Colpitts TL, Friedman PN;
 PI Gordon J, Granadosen, Hodges SC, Klass MR, Kratochvil JD;
 PI Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;

DR WPI: 1999-045802/04.

PT New purified polynucleotide BL172 derivatives and encoded
 PT polypeptides - useful in the identification of markers, which are
 PT indicative of urinary tract diseases or conditions

PS Claim 1; Page 89; 114pp; English.

CC Sequences AA01683-85 represent overlapping clones specific for urinary
 CC tract tissue library BL172. The BL172 polynucleotides are used to detect
 CC target BL172 polynucleotides and BL172 mRNA and the BL172 polypeptides
 CC are used to detect BL172 antigens/antibodies in a test sample. The
 CC identification of certain markers in these methods are indicative of the
 CC presence of urinary tract disease, especially cancer. Additionally, the
 CC polypeptides are used to detect antibodies that bind specifically to a
 CC BL172 epitope derived from a BL172 polypeptide. Host cells containing a
 CC recombinant expression vector comprising the BL172 polynucleotide
 CC sequences are used to produce the polypeptides containing BL172
 CC epitopes. The methods aid in diagnosis, detection, staging, monitoring,
 CC prognostication, in vivo imaging, prevention and treatment of diseases
 CC or conditions associated with BL172, especially urinary tract cancer.

XX Sequence 236 BP; 49 A; 69 C; 62 G; 56 T; 0 other;

Query Match 30.9%; Score 236; DB 20; Length 236;
 Best Local Similarity 100.0%; Pred. No. 1e-59;
 Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 TTTCACGACCCCTGCAGCACTGTTGCTATGATGATCCGCGTGGCCCTGGCCAGGACC 221
 |||||
 Db 1 TTTCACGACCCCTGCAGCACTGTTGCTATGATGATCCGCTGGCCCTGGCCAGGACC 60

QY 222 CAGACGTGTGGAACACCTGACCTTCAGATCTGCTTTGAGCAGTGCCTCCCTGGACCTTC 281
 |||||
 Db 61 CAGACGTGTGGAACACCTGACCTTCAGATCTGCTTTGAGCAGTGCCTCCCTGGACCTTC 120

QY 282 ATGCTGAAGCTGATAAACCGAAGCTGCACTGACCCCGGACCTCGGATGACAGGCTTTGT 341
 |||||
 Db 121 ATGCTGAAGCTGATAAACCGAAGCTGCACTGACCCCGGACCTCGGATGACAGGCTTTGT 180

QY 342 CGCAGTGTCAAGCTATGAGACATCAGGGGAGAGTACCTCTGGATCTCTCTTCT 397
 |||||
 Db 181 CGCAGTGTCAAGCTATGAGACATCAGGGGAGAGTACCTCTGGATCTCTCTTCT 236

RESULT 13

AAx01683
 ID AAX01683 standard; DNA; 196 BP.

AC AAX01683;

DT 21-APR-1999 (first entry)

DE Urinary tract tissue library BL172 gene specific cDNA clone 1554838.

KW BL172; urinary tract; marker; cancer; recombinant; human; ss.

OS Homo sapiens.

PN M09855656-A1.

PD 10-DEC-1998.

PF 05-JUN-1998; 98MO-US11693.

PR 05-JUN-1997; 97US-0869579.

PA (ABBO) ABBOTT LAB.

PI Billing-medel PA, Cohen M, Colpitts TL, Friedman PN;

PI Gordon J, Granadosen, Hodges SC, Klass MR, Kratochvil JD;

PI Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;

DR WPI; 1999-045802/04.

PT New purified polynucleotide BL172 derivatives and encoded

PT polypeptides - useful in the identification of markers, which are

PT indicative of urinary tract diseases or conditions

PS Claim 1; Page 89; 114pp; English.

CC Sequences AAX01683-85 represent overlapping clones specific for urinary
 CC tract tissue library BL172. The BL172 polynucleotides are used to detect
 CC target BL172 polynucleotides and BL172 mRNA and the BL172 polypeptides
 CC are used to detect BL172 antigens/antibodies in a test sample. The
 CC identification of certain markers in these methods are indicative of the
 CC presence of urinary tract disease, especially cancer. Additionally, the
 CC polypeptides are used to detect antibodies that bind specifically to a
 CC BL172 epitope derived from a BL172 polypeptide. Host cells containing a
 CC recombinant expression vector comprising the BL172 polynucleotide
 CC sequences are used to produce the polypeptides containing BL172
 CC epitopes. The methods aid in diagnosis, detection, staging, monitoring,
 CC prognostication, in vivo imaging, prevention and treatment of diseases
 CC or conditions associated with BL172, especially urinary tract cancer.

CC Sequence 196 BP; 38 A; 70 C; 44 G; 44 T; 0 other;

Query Match 25.7%; Score 196; DB 20; Length 196;
 Best Local Similarity 100.0%; Pred. No. 6.1e-48;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCACAGCAACCAACCCAGAGCCATGGCTCCCGAGGCTGCATGCTGCTTTGCCA 60
 |||||
 Db 1 CTCACAGCAACCAACCCAGAGCCATGGCTCCCGAGGCTGCATGCTGCTTTGCCA 60

QY 61 TTTTCTGCATCTCCAGGCTCCTCTGCTCAGACGAGCCCGAGTGGCCCATGACTCCTT 120
 |||||
 Db 61 TTTTCTGCATCTCCAGGCTCCTCTGCTCAGACGAGCCCGAGTGGCCCATGACTCCTT 120

QY 121 ACCTGATCTGTGCGCACACACAGAGATGTGGGCAAGTTCTACAGCCCTGCGAGC 180
 |||||
 Db 121 ACCTGATCTGTGCGCACACACAGAGATGTGGGCAAGTTCTACAGCCCTGCGAGC 180

QY 181 ACTGTGCTATGATGA 196
 |||||
 Db 181 ACTGTGCTATGATGA 196

RESULT 14

AAS92056
 ID AAS92056 standard; cDNA; 1338 BP.

AC AAS92056;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #27860.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN M0200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001MO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR P-PDSB; ABG27869.

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics; forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity

PS Claim 1; SEQ ID NO 27860; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 23:07:27 : Search time 1692 Seconds
(without alignments)
6086.393 Million cell updates/sec

Title: US-09-092-297-5

Perfect score: 763
Sequence: 1 CTCGACTGCACACACCCAGA.....ATTAATTATGACTTTATA 763

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: em_estba:*
2: em_esthum:*
3: em_estln:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vitc:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	628.8	82.4	729	10	BG775668 602650023
C 2	617.4	80.9	729	10	BG775059 602650023
C 3	483.8	63.4	533	9	AM991605 R01-BN000
C 4	436.6	57.2	475	9	AA195677 2732104.s
C 5	416.6	54.6	417	9	AA455904 aa14e02.s
C 6	401.5	52.6	412	9	AA456370 aa14e02.s
C 7	381.8	50.0	604	10	BG680336 602629230
C 8	313.6	41.1	354	9	AM265494 xg70c11.x
C 9	296.4	38.8	308	9	AM265432 xg52h11.x
C 10	251.4	32.9	253	10	BF911380 IL2-UT007
C 11	208.6	27.3	245	10	BF913866 IL3-UT011
C 12	207	27.1	245	10	BF914311 IL3-UT011
C 13	189.2	24.8	554	10	BE664072 148225 MA
C 14	186	24.4	555	10	BE664075 148231 MA
C 15	174.8	22.9	201	10	BF911379 IL2-UT007
C 16	90.6	11.9	366	9	BE122850 02_15 Hum
C 17	86	11.3	275	9	BE122850 02_15 Hum
					AT572169 tes3e06.x

18	80.2	10.5	537	9	AA024389
C 19	64.8	8.5	481	9	AI806131
C 20	62	8.1	103	10	BG951874
C 21	58.6	7.7	464	9	AI275379
C 22	57.8	7.6	455	9	AA454642
C 23	57.8	7.6	523	9	AA456254
C 24	57	7.5	454	9	AI081371
C 25	49.2	6.4	130	9	AW795497
C 26	42.6	5.6	834	10	BM010458
C 27	41.2	5.4	575	9	AA760247
C 28	40.4	5.3	844	12	CNS0052P
C 29	39.6	5.2	717	10	BG122553
C 30	39.6	5.2	749	10	BF128901
C 31	39.6	5.2	794	10	BE300864
C 32	39.6	5.2	801	10	BG330279
C 33	39.6	5.2	853	10	CNS03498
C 34	39.4	5.2	815	12	CNS03498
C 35	39.2	5.1	934	10	BE567738
C 36	38.8	5.1	914	10	BE539747
C 37	38.8	5.1	1101	12	CNS00DJC
C 38	38.6	5.1	611	9	AA474076
C 39	38.4	5.0	1003	10	BG331909
C 40	38.2	5.0	283	10	H41708
C 41	38	5.0	424	9	AM382176
C 42	38	5.0	688	10	BG975448
C 43	37.6	4.9	869	10	BE685231
C 44	37.6	4.9	585	12	A2100949
C 45	37.6	4.9	914	10	BE903336

ALIGNMENTS

RESULT 1
BG775668/c
LOCUS
DEFINITION
602650023T1 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:4761029 3'
mRNA sequence.
ACCESSION
BG775668
VERSION
BG775668
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 828)
NIH-MGC <http://mgs.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LCM1612 row: n column: 06
High quality sequence start: 16
High quality sequence stop: 806.
Location/Qualifiers
1. 828

FEATURES
Source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4761029"
/clone_lib="NIH_MGC_40"
/tissue_type="Carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. directionally cloned sites XhoI sites using the following 5' adaptor: GGCCAGAG(G). Library constructed by

RESULT 2	BG775059	729 bp	mRNA	linear	EST 15-MAY-2001
LOCUS	BG775059				
DEFINITION	60265002.F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4761029 5', mRNA sequence.				
ACCESSION	BG775059				
VERSION	BG775059.1	GI:14045376			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 729)				
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D.				

Query Match	80.9%	Score 617.4	DB 10	Length 729
Best Local Similarity	98.8%	Pred. 2.1e-161		
Matches 673	Conservative	1	Mismatches 2	Indels 5
			Gaps	5
QY 61	TTTTGCACATTCACAGCTCCTCTGCTACACACGAGCCCCAGTAGGCCCCCATGATCTT	120		
Db 2	TTTTGCATCTCCAGGCTCCTCTGCTACACAGGAGGCCCATGATCTT	61		
QY 121	ACCTGATGCTGTGCCAGCCACACACAGATGTGGGGACAAGTCTACAGCCCTCGAGC	180		
Db 62	ACCTGATGCTGTGCCAGCCACACAGATGTGGGGACAAGTCTACAGCCCTCGAGC	121		
QY 181	ACGTGTCTATGATGATGGCGGTGCGCCCTTGGCCAGAG-CCAGACGTGTGAAATGC	239		
Db 122	ACGTGTCTATGATGATGGCGGTGCGCCCTTGGCCAGAGCCCGAGCGTGTGAAATGC	181		
QY 240	ACCTTCAGAGTCTGCTTTGAGCAGTGTGCCCTTGACCTTCATGTGTGAAGCTATAAC	299		
Db 182	ACCTTCAGAGTCTGCTTTGAGCAGTGTGCCCTTGAGACCTTCATGTGTGAAGCTATAAC	241		
QY 300	CAGAACTGCGACTAGCCCCGGAACCTGGATACAGGGCTTGTCCAGTGTAGCTAATGG	359		
Db 242	CAGAACTGCGACTAGCCCCGGAACCTGGATACAGGCTTGTCCAGTGTAGCTAATGG	301		
QY 360	AACATCAGGGGAACGATGACTCTCGATTCCTCTCGGGTGGGCTTGGAAAGAGCC	419		
Db 302	AACATCAGGGGAACGATGACTCTCGATTCCTCTCGGGTGGGCTTGGAAAGAGCC	361		
QY 420	TGTGTTCACCTGAGATCTGGGATGTGTAGTGGCTTTGGGGGCCAGAGAAACACACT	479		
Db 362	TGTGTTCACCTGAGATCTGGGATGTGTAGTGGCTTTGGGGGCCAGAGAAACACACT	421		
QY 480	CAACTGGCCACTCTATTC-TGTGACCTATGTGAGGCCACACCCTGCGGCTGCCCTGAGAG	538		
Db 422	CAACTGGCCACTCTATTC-TGTGACCTATGTGAGGCCACACCCTGCGGCTGCCCTGAGAG	481		
QY 539	GCCACAGATCCCTTCATGAATTTGTGACACATGAGATGGTGTGCTGATGGGGGCC	598		
Db 482	GCCACAGATCCCTTCATGAATTTGTGACACATGAGATGGTGTGCTGATGGGGGCC	541		
QY 599	AGGAGACTCGAACCTCTCGATGAGCCCATATGGCCAACTCAACCCGGCACCCCAAG	658		
Db 542	A-GGACTCGAACCTCTCGATGAGCCCATATGGCCAACTCAACCCGGCACCCCAAG	600		

QY 659 GCTGGCTGGGGAACCTTCA-CCCTTCTGTGAGATTTTCATCATCA-AGTTCCTTC 716
 |||||||
 Db 601 GCTGGCTGGGGAACCTTCA-CCCTTCTGTGAGATTTTCATCATCA-AGTTCCTTC 660
 |||||||
 QY 717 TATCCAGAGCAAGACACAGG 737
 |||||||
 Db 661 TATCCAGAGCAAGACACAGG 681
 |||||||

RESULT 3
 AM991605 533 bp mRNA linear EST 05-JUN-2000
 LOCUS RCI-BN0005-030500-015-g07 BN0005 Homo sapiens cDNA, mRNA sequence.
 DEFINITION AM991605
 ACCESSION AM991605.1 GI:8251660
 VERSION EST.
 KEYWORDS human.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS 1 (bases 1 to 533)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brlones, M.R.,
 Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 2020263
 CONTACT: Simpson A.J.G.
 LABORATORY OF Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?rl=6t2-RC1-BN0005-030
 500-015-g07&t3=2000-05-03&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 16
 High quality sequence stop: 533.
 Location/Qualifiers
 1..533
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="BN0005"
 /dev_stage="Adult"
 /note="Organ: breast normal; Vector: puc18; Site:1; SmaI;
 Site:2; SmaI: A mini-library was made by cloning products
 derived from ORSTES PCR (U.S. Letters Patent application
 No.196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 116 a 156 c 150 g 111 t

Query Match 63.48; Score 483.8; DB 9; Length 533;
 Best Local Similarity 98.7%; Pval. NO. 3.4e-124;
 Matches 518; Conservative 1; Mismatches 3; Indels 3; Gaps 3;

QY 135 CAGCCACACAGAGATGCGGGGAGAAAGTTTCACACACCCCGCAGACATGTTGCTATAT 194
 |||||||
 Db 12 CAGCCACACAGAGATGCGGGGAGAAAGTTTCACACACCCCGCAGACATGTTGCTATAT 68
 |||||||
 QY 195 GATGCCCTGTGCTTGGCCAGACAGACAGATGTGAAACTGCACCTCAGAGTTCGC 254
 |||||||
 Db 69 GATGCCCTGTGCTTGGCCAGACAGACAGATGTGAAACTGCACCTCAGAGTTCGC 128
 |||||||

QY 255 TTGAGACAGTGCCTGCCCCCTGGAGACCTTCATAGTGAAGCTGTAACCAAGACAGTGCAGTCA 314
 |||||||
 Db 129 TTGAGACAGTGCCTGCCCCCTGGAGACCTTCATAGTGAAGCTGTAACCAAGACAGTGCAGTCA 188
 |||||||
 QY 315 GCCCGGACCTCGGATGACAGAGCTTTGTGCGAGTGTACGCTAATGAAATCAAGGGAGAG 374
 |||||||
 Db 189 GCCCGGACCTCGGATGACAGAGCTTTGTGCGAGTGTACGCTAATGAAATCAAGGGAGAG 248
 |||||||

QY 375 ATGACTCCTGATTCCTCTCCCTGGGAGGCGGAGAAAGAGGCTGTACCTGAGA 434
 |||||||
 Db 249 ATGACTCCTGATTCCTCTCCCTGGGAGGCGGAGAAAGAGGCTGTACCTGAGA 308
 |||||||

QY 435 TCTGGAGTGTGATGCTGCTTTTGGGGCCAGAGAAACACACTCAACTGCCACTTCA 494
 |||||||
 Db 309 TCTGGAGTGTGATGCTGCTTTTGGGGCCAGAGAAACACACTCAACTGCCACTTCA 368
 |||||||

QY 495 TTCTGACCTGTCTAGGCCACCTGCCGCTGCCCTGAGAGGCCACAGTCCCTT 554
 |||||||
 Db 369 TTCTGACCTGTCTAGGCCACCTGCCGCTGCCCTGAGAGGCCACAGTCCCTT 428
 |||||||

QY 555 CTGAAATCTGGACACATGAGATGCTGCTGATGAGGGGCCAGGAGCTGTGAACCT 614
 |||||||
 Db 429 CTGAAATCTGGACACATGAGATGCTGCTGATGAGGGGCCAGGAGCTGTGAACCT 488
 |||||||

QY 615 CCGATGACCCCTATGCGCCACATCAACCCGACACCCCAAG 659
 |||||||
 Db 489 CCGATGACCCCTATGCGCCACATCAACCCGACACCCCAAG 533
 |||||||

RESULT 4
 AA195677 475 bp mRNA linear EST 19-MAY-1997
 LOCUS zrc3h04.s1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:665143
 DEFINITION AA195677
 ACCESSION AA195677.1 GI:1785355
 VERSION EST.
 KEYWORDS human.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS 1 (bases 1 to 475)
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
 Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins,
 M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore,
 B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
 Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisan, E.,
 Underwood, K., Wohlmann, E., Waterston, R., Wilson, R. and Marra, M.
 Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)
 97044478
 CONTACT: Wilson-RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estevenson.wustl.edu
 This clone is available royalty-free through LNC; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 857 Std Error: 0.00
 High quality sequence stop: 308.
 Location/Qualifiers
 1..475
 /organism="Homo sapiens"
 /db_xref="GDB:5427175"
 /db_xref="taxon:9606"
 /clone="IMAGE:665143"
 /clone_id="Soares_NHMPu_S1"
 /tissue_type="Pooled human melanocyte, fetal heart, and
 pregnant uterus"
 /lab_host="DH10B"
 /note="Organ: mixed (see below); Vector: pT73D-Pac

Query Match	57.2%	Score 436.6	DB 9	Length 475
Best Local Similarity	98.9%	Pred. No. 4,8e-111		
Matches 469	Conservative 1	Mismatches 1	Indels 3	Gaps 3
QY 291	CTGATTAACACGAACCTCGGACTACAGCCCGGACCTCGATGACAGGCTTTGTGCGAGTGTG			350
Db 475	CTGATTAACACGAANTGCGACTACAGCCCGGACCTCGATGACAGGCTTTGTGCGAGTGTG			416
QY 351	AG-CTAATGGAACATCAGGGGAGAGATGATCTCTTGATTTCTCTTC-CTGGGTGGCGCTG			408
Db 415	AGCCTTAATGGAACATCAGGGGAGAGATGATCTCTTGATTTCTCTTCCTTCCTGCTGCGCTG			356
QY 409	GAGAAAGAGGCTGTGTTACTGATGATGTGGGATGGCTAGTGGC-TGTTTGGGGGCCAGA			467
Db 355	GAGAAAGAGGCTGTGTTACTGATGATGTGGGATGGCTAGTGGCTGTTTGGGGGCCAGA			296
QY 468	GAACACACACTCACTGCTCCACTTCATTCGTGTGACCTGTGTGAGGCCACACCTGCGCT			527
Db 295	GAACACACACTCACTGCTCCACTTCATTCGTGTGACCTGTGTGAGGCCACACCTGCGCT			236
QY 528	GCCCTGAGGAGGCCACAGTCCCTTTTAAATTTCTGGACAGCATGATGATGCTGTGCT			587
Db 235	GCCCTGAGGAGGCCACAGTCCCTTTTAAATTTCTGGACAGCATGATGATGCTGTGCT			176
QY 588	GATGGGGGCCAGGAGACTGTAACCCCTGATGACCCCTATGGGCAACATCAACCCGGC			647
Db 175	GATGGGGGCCAGGAGACTGTAACCCCTGATGACCCCTATGGGCAACATCAACCCGGC			116
QY 648	ACCAACCCCAAGGCTGTGGTGGGGAACCTTTCACCTTCTGTGAGATTTTCCATCATCTGAA			707
Db 115	ACCAACCCCAAGGCTGTGGTGGGGAACCTTTCACCTTCTGTGAGATTTTCCATCATCTGAA			56
QY 708	GTCTCTTCTTCTCCAGGAGCAAGACACAGATCATTAATTAATTAATGATGACTTTA			761
Db 55	GTCTCTTCTTCTCCAGGAGCAAGACACAGATCATTAATTAATTAATGATGACTTTA			2

RESULT 5
AA455904/c

LOCUS AA455904 417 bp mRNA linear EST 06-JUN-1997

DEFINITION aa14602.s1 Soares_NHNPu_S1 Homo sapiens cDNA clone IMAGE:813242

ACCESSION 3, mRNA sequence.
AA455904

VERSION AA455904.1 GI:2178680

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryote: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.

TITLE Mashu-Merck EST Project 1997

JOURNAL Unpublished (1997)

COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800

```

FEATURES
source
location/Qualifiers
1..417
/organism="Homo sapiens"
/db.xref="GDB:6044152"
/db.xref="taxon:9606"
/clone_image="IMAGE:813242"
/clone_id="Scases_NbHNPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10b"
/note="Organ: mixed (see below); Vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site:2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbM, pregnant uterus NbHPU, and fetal heart NbHNPu), were mixed, and ss clones were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."
BASE COUNT
99 a 108 c 117 g 93 t
ORIGIN

```

Query Match	54.6%	Score 416.6	DB 9	Length 417
Best Local Similarity	99.8%	Pred. No. 1.7e-105		
Matches 416	Conservative 1	Mismatches 0	Indels 0	Gaps 0
QY 346	GTGTACGCTAAATGGAACATCAGGGGAGACGATGACTCTCGATTTCTCTTCCCTGGGTGGGC	405		
Db 417	GTGTACGCTAATGGAACATCAGGGGAGACGATGACTCTCGATTTCTCTTCCCTGGGTGGGC	358		
QY 406	CTGGAGAAAGAGGCTGGTGTCTTACTCGAATCTGGGATCTCTGAGTGGCTGTTGGGGCCA	465		
Db 357	CTGGAGAAAGAGGCTGGTGTCTTACTCGAATCTGGGATCTGAGTGGCTGTTGGGGGCCA	298		
QY 466	GAGAAACACACACTCACTGCGCCACTTATCTGTGACCTGTGAGGCCCCACCCCTGGCG	525		
Db 297	GAGAAACACACACTCACTGCGCCACTTATCTGTGACCTGTGAGGCCCCACCCCTGGCG	238		
QY 526	CTGCCCTGTAGAGAGGCCACAGGTCCCTCTAGAAATTTGTGACACATGAGATGGCGTNG	585		
Db 237	CTGCCCTGTAGAGAGGCCACAGGTCCCTCTAGAAATTTGTGACACATGAGATGGCGTNG	178		
QY 586	CTGATGTGGGGGCCCAAGGAGCTCTGAACCTCTCTGATGAGACCCSTATGGCCAATCAACCG	645		
Db 177	CTGATGTGGGGGCCCAAGGAGCTCTGAACCTCTCTGATGAGACCCSTATGGCCAATCAACCG	118		
QY 646	GCACACACCCCAAGGCTGAGCTGGGAGAACCTTCACCCCTCTGTGAGATTTTCATCATCTC	705		
Db 117	GCACACACCCCAAGGCTGAGCTGGGAGAACCTTCACCCCTCTGTGAGATTTTCATCATCTC	58		
QY 706	AAGTCTCTTCTATTCACGAGCAAAAGCAGAGATCATATAAATTTATGACTTTAT	762		
Db 57	AAGTCTCTTCTATTCACGAGCAAAAGCAGAGATCATATAAATTTATGACTTTAT	1		
RESULT 6				
LOCUS	AA456370			
DEFINITION	AA456370	412 bp	mRNA	linear
ACCESSION	aa14602.r1	Soares.NHMPU.SI	Homo sapiens	CDNA clone IMAGE:813242
VERSION	AA456370			
KEYWORDS	EST.	GI:2178946		
SOURCE	human.			
ORGANISM	Homo sapiens			

RESULT 8
AM265494 354 bp mRNA linear EST 28-DEC-1999
LOCUS xg70c11.x1 NCI_CGAP_HN7 Homo sapiens cDNA clone IMAGE:2755988 3'
DEFINITION mRNA sequence.
ACCESSION AM265494
VERSION AM265494.1 GI:6642310
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 354)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Edward Shillitoe Ph.D., Silvio Gutkind Ph.D.,
Chichanok Leethanakul D.D.S., Michael Emmert-Buck M.D. Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html

FEATURES
source
Location/Qualifiers
1..354
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2755988"
/clone_1lb="NCI_CGAP_HN7"
/tissue_type="normal squamous epithelium, floor of mouth"
/lab_host="DH10B"
/note="Vector: PAMP10; cDNA made by oligo-dT priming.
Non-directionally cloned into the UDG sites of PAMP10.
Size-selected on agarose gel, average insert size 500 bp.
Primary library; non-amplified. cDNA library
Preparation: David B. Krizman, Ph.D (NCI). Reference:
Krizman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT 99 a 102 c 78 g 75 t

ORIGIN
Query Match 41.1%; Score 313.6; DB 9; Length 354;
Best Local Similarity 99.7%; Pred. No. 8.9e-77;
Matches 313; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 450 GCGCTTTGGGGGCGCAGGAACAACACACTGACCTGACCTGACCTGCT 509
DB 9 GCGCTTTGGGGGCGCAGGAACAACACACTGACCTGACCTGACCTGCT 68
QY 510 GAGGCCACCTGCGCGCTGCGCTGAGAGGCCACAGGTCCTCTGAGAAATTCGACA 569
DB 69 GAGGCCACCTGCGCGCTGCGCTGAGAGGCCACAGGTCCTCTGAGAAATTCGACA 128
QY 570 GCATGAGATGCGTGTGCTGATGAGGGGCCAGGACTGAAACCTCTGATGAGACCCCTAT 629
DB 129 GCATGAGATGCGTGTGCTGATGAGGGGCCAGGACTGAAACCTCTGATGAGACCCCTAT 188
QY 630 GGCACAATCAACCCGGGACACACCCCAAGGCTGGGGAACCTTCACCTTCTGTGA 689
DB 189 GGCACAATCAACCCGGGACACACCCCAAGGCTGGGGAACCTTCACCTTCTGTGA 248
QY 690 GATTTTCATCATCTCAAGTTCTCTTCTATCCAGAGCAAAAGCAGATCATTAATAAT 749

RESULT 9
AM265432 308 bp mRNA linear EST 28-DEC-1999
LOCUS xg52h11.x1 NCI_CGAP_HN7 Homo sapiens cDNA clone IMAGE:2754309 3'
DEFINITION mRNA sequence.
ACCESSION AM265432
VERSION AM265432.1 GI:6642248
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 308)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Edward Shillitoe Ph.D., Silvio Gutkind Ph.D.,
Chichanok Leethanakul D.D.S., Michael Emmert-Buck M.D. Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html

FEATURES
source
Location/Qualifiers
1..308
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2754309"
/clone_1lb="NCI_CGAP_HN7"
/tissue_type="normal squamous epithelium, floor of mouth"
/lab_host="DH10B"
/note="Vector: PAMP10; cDNA made by oligo-dT priming.
Non-directionally cloned into the UDG sites of PAMP10.
Size-selected on agarose gel, average insert size 500 bp.
Primary library; non-amplified. cDNA library
Preparation: David B. Krizman, Ph.D (NCI). Reference:
Krizman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT 69 a 97 c 73 g 69 t

ORIGIN
Query Match 38.8%; Score 296.4; DB 9; Length 308;
Best Local Similarity 99.0%; Pred. No. 5.3e-72;
Matches 297; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 450 GCGCTTTGGGGGCGCAGGAACAACACACTGACCTGACCTGACCTGCT 509
DB 9 GCGCTTTGGGGGCGCAGGAACAACACACTGACCTGACCTGACCTGCT 68
QY 510 GAGGCCACCTGCGCGCTGCGCTGAGAGGCCACAGGTCCTCTGAGAAATTCGACA 569
DB 69 GAGGCCACCTGCGCGCTGCGCTGAGAGGCCACAGGTCCTCTGAGAAATTCGACA 128
QY 570 GCATGAGATGCGTGTGCTGATGAGGGGCCAGGACTGAAACCTCTGATGAGACCCCTAT 629
DB 129 GCATGAGATGCGTGTGCTGATGAGGGGCCAGGACTGAAACCTCTGATGAGACCCCTAT 188
QY 630 GGCACAATCAACCCGGGACACACCCCAAGGCTGGGGAACCTTCACCTTCTGTGA 689

Db 189 GGGCAACATCAACCGGACCCCAAGGCTGGGGAACCTTACCCCTTGTGA 248
QY 650 GATTTCATCATCTCAAGTCTCTTATCCAGAGCAACAGATCATTAAT 749
Db 249 GATTTCATCATCTCAAGTCTCTTATCCAGAGCAACAGATCATTAAT 308

RESULT 10
LOCUS BF911380 253 bp mRNA linear EST 18-JAN-2001
DEFINITION IL2-UT0073-081100-213-D02.1 UT0073 Homo sapiens CDNA, mRNA
ACCESSION BF911380
VERSION BF911380.1 GI:12302838
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 253)
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL2&t2=IL2-UT0073-
081100-213-D02.1&t3=2000-11-08&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 220.
Location/Qualifiers
1. 253
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_11b="UT0073"
/dev_stage="Adult"
/note="Organ: uterus; tumor; Vector: puc18; site_1: SmaI;
site_2: SmaI; A mini-library was made by cloning products
derived from ORESFES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 51 a 67 c 79 g 56 t
ORIGIN
Query Match 32.9%; Score 251.4; DB 10; Length 253;
Best Local Similarity 99.6%; Pred. No. 1.8e-59;
Matches 252; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 358 GGAATCATGAGGAGAGTCTCTGATCTCTCTCTGCTGGGCTGGAGAAAG 417
Db 1 GGAATCATGAGGAGAGTCTCTGATCTCTCTCTGCTGGGCTGGAGAAAG 60
QY 418 GCTGCTGTTACTGAGATCTGGAGTCTGAGTCTGTTGGGGGCAAGAAACACA 477
Db 61 GCTGCTGTTACTGAGATCTGGAGTCTGAGTCTGTTGGGGGCAAGAAACACA 120

QY 478 CTCACCTGCCACTTATCTGTGACCTGTCTGAGGCCCCACCCCTGCTGAGCA 537
Db 121 CTCACCTGCCACTTATCTGTGACCTGTCTGAGGCCCCACCCCTGCTGAGCA 180
QY 538 GGGCCACAGTCCCTCTAGAAATCTGGACGATGAGTCCGCTGATGGGGCC 597
Db 181 GGGCCACAGTCCCTCTAGAAATCTGGACGATGAGTCCGCTGATGGGGCC 240
QY 598 CAGGACTCTGAA 610
Db 241 CAGGACTCTGAA 253

RESULT 11
LOCUS BF913866/c 245 bp mRNA linear EST 18-JAN-2001
DEFINITION IL3-UT0114-301100-357-C03 UT0114 Homo sapiens CDNA, mRNA sequence.
ACCESSION BF913866
VERSION BF913866.1 GI:12305324
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 245)
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&t2=IL3-UT0114-
301100-357-C03&t3=2000-11-30&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 245.
Location/Qualifiers
1. 245
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_11b="UT0114"
/dev_stage="Adult"
/note="Organ: uterus; tumor; Vector: puc18; site_1: SmaI;
site_2: SmaI; A mini-library was made by cloning products
derived from ORESFES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 55 a 63 c 82 g 45 t
ORIGIN
Query Match 27.3%; Score 208.6; DB 10; Length 245;
Best Local Similarity 95.1%; Pred. No. 1.5e-47;
Matches 214; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 494 ATTCTGACCTGTCTGAGAGCCACCTCCGCTGCTGAGAGGCGCCACAGTCCCT 553
Db 225 ATTCTGACCTGTCTGAGAGGAGCCACCTCCGCTGCTGAGAGGCGCCACAGTCCCT 166

QY	554	TCGAAATTTCTGGACAGATGGATGGCGGTCTGATGGGGCCACAGGACTCTGAACC	613
QY	165	TCCTAAATTTCTGGACAGATGGATGGCGGTCTGATGGGGCCACAGGACTCTGAACC	106
QY	614	TCCTGATGACCCCTATGGCAACATCAACCCGGGCAACCCCAAGGCTGGGGGAACC	673
DB	105	TCCTGATGACCCCTATGGCAACATCAACCCGGGCAACCCCAAGGCTGGGGGAACC	46
QY	674	CTTCACCCCTCTGTGAGATTTCCATCAATCTCAATGTTCTCTCTCA 718	
DB	45	CTTCACCCCTCTGTGAGATTTCCATCAATCTCAATGTTCTCTCTCA 1	

RESULT	12	
BF914311/c		
LOCUS	BF914311	245 bp mRNA linear EST 18-JAN-2001
DEFINITION	IL3-U70114-301100-356-E11 U70114 Homo sapiens CDNA, mRNA sequence.	
ACCESSION	BF914311	
VERSION	BF914311.1	GI:12305769
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS

Eulazeyota, Metazoa; Chordata; Craniata; Vertebrati; Euteleostomi; Mammalia; Eutheria; Primates; Carnivorni; Hominiidae; Homo.
1. (bases 1 to 245)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.,

TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
COMMENT	Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel.: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?c1=113&c2=IL3-UT0114>
301100-356-Effect3-2000-11-30&c4=1)
Seq primer: puc 18 forward
High quality sequence stop: 245.

FEATURES	SOURCE	BASE COUNT	ORIGIN
Location/Qualifiers	1. .245		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_lib="U70114"		
	/dev_stage="Adult"		
	/note="Organ: uterus; tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."		
		55 a	63 c 81 g 46 t

Query Match	27.1%	Score 207	DB 10	Length 245
Best Local Similarity	94.7%	Pred. No.	4,2e+47	
Matches 213;	Conservative 1;	Mismatches 11;	Indels 0;	Gaps 0.

Oy	494	ATTCTGTGACCTGTCTGAGGCCACCCTGCCTCCCTCCTGAGAGGCCACAGTCCCT	553
Db	225	ATTCGGGCTCTCTGGGAGGCCACACTCCCTGCCTGAGAGGCCACAGTCCCT	166

QY	554	TCTGGAATCTGGGACAGATGAGAGGCGCTGCTATGCGGGGCCAGGAGACTGTGAACC	613
Db	165	TCTGGAATCTGGGACAGATGAGAGGCGCTGCTATGCGGGGCCAGGAGACTGTGAACC	106
QY	614	TCTGATGATACCCCTGTATGGCCAAATCAACCCGGCACCCCAAGGCTGGGGGAACC	673
Db	105	TCTGATGATACCCCTGTATGGCCAAATCAACCCGGCACCCCAAGGCTGGATGGGGAACC	46
QY	674	CTTCAACCCCTCTGTGAGATTTCCATCAATCAATCTCAATTTCTCTCTCA	718
Db	45	CTTCAACCCCTCTGTGAGATTTCCATCAATCTCAATTTCTCTCTCA 1	

RESULT	13
BEG64072	
LOCUS	BEG64072
DEFINITION	188225 MARC 4BOV Bos taurus CDNA 5' , mRNA linear EST 25-APR-2001
ACCESSION	BEG64072
VERSION	BEG64072.1 GI:10023165
KEYWORDS	EST.
SOURCE	COW,
ORGANISM	Bos taurus

Euarchyotia, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos.

TITLE	Sequence evaluation of four pooled-tissue normalised bovine cDNA libraries and construction of a gene index for cattle
JOURNAL	Genome Res. 11 (4), 626-630 (2001)
MEDLINE	21180013
COMMENT	Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4356
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 1
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCATTATGACCATT
BACKWARD: GTTTCACAGTCACGACG
Plate: 51 row: M column: 8
Seq primer: ATTAGGTGACACTATTAG.

FEATURES	
Source	Location/Qualifiers
	1..554
	/organism="Bos taurus"
	/db_xref="taxon:9913"
	/clone_lib="MARC 4BOV"
	/tissue_type="pooled"
	/lab_host="DH10B"
	/note="Vector: pCMV SPORT6; Site 1: XbaI; Site 2: XhoI;
	library made from pooled tissue from day 20 and day 40
	embryos."
BASE COUNT	128 a 154 c 138 g 134 t
ORIGIN	
Query Match	24.8% Score 189.2; DB 10; Length 554;
Best Local Similarity	67.1% Pred. No. 3.7e+42;

	Matches	294; Conservative	0; Mismatches	123; Indels	21; Gaps
QY	12	CCACCAGAGCATGGCTCCCGAGAGCGCATCGAGCTGCTTGGCCATTTTTCGACATC	71		
Db	6	CAATCAAGAGCCATATCTACACGACGAGCTGCATCTGGGCTGTGGCCCTCTCTGCATCC	65		
QY	72	TCCAGGCTCTCTGTCACACGAGAGCCCGAGTGGGCCCATCTACTCTTACCTGATGATCG	131		
Db	65	CTTTTGGCTTTCTTTTTCATATCTGCTCCCTCCAGTCTCCCTCCACAGAGATAGATGTTGGCTG	125		

REFERENCE	TITLE	JOURNAL MEDLINE COMMENT
132	132	132
126	126	126
192	192	192
186	186	186
252	252	252
246	246	246
291	291	291
306	306	306
351	351	351
366	366	366
411	411	411
426	426	426
LOCUS	LOCUS	LOCUS
DEFINITION	DEFINITION	DEFINITION
ACCESSION	ACCESSION	ACCESSION
VERSION	VERSION	VERSION
KEYWORDS	KEYWORDS	KEYWORDS
SOURCE	SOURCE	SOURCE
ORGANISM	ORGANISM	ORGANISM
BOV	BOV	BOV
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; 1 (bases 1 to 555)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; 1 (bases 1 to 555)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; 1 (bases 1 to 555)
Smith, T.P.L., Groves, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G., Pettea, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and Keele, J.W.	Smith, T.P.L., Groves, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G., Pettea, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and Keele, J.W.	Smith, T.P.L., Groves, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G., Pettea, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and Keele, J.W.
Sequence evaluation of four pooled tissue normalized bovine cDNA libraries and construction of a gene index for cattle	Sequence evaluation of four pooled tissue normalized bovine cDNA libraries and construction of a gene index for cattle	Sequence evaluation of four pooled tissue normalized bovine cDNA libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)	Genome Res. 11 (4), 626-630 (2001)	Genome Res. 11 (4), 626-630 (2001)
Contact: Smith TPL	Contact: Smith TPL	Contact: Smith TPL
USA, ARS, US Meat Animal Research Center	USA, ARS, US Meat Animal Research Center	USA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA	PO Box 166, Clay Center, NE 68933-0166, USA	PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366	Tel: 402 762 4366	Tel: 402 762 4366
Fax: 402 762 4390	Fax: 402 762 4390	Fax: 402 762 4390
Email: smith@email.marc.usda.gov	Email: smith@email.marc.usda.gov	Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.	Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.	Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.
PCR Primers	PCR Primers	PCR Primers
FORWARD: AGGAACACGCTATGACCAT	FORWARD: AGGAACACGCTATGACCAT	FORWARD: AGGAACACGCTATGACCAT
BACKWARD: GTTTCACGTCACGACG	BACKWARD: GTTTCACGTCACGACG	BACKWARD: GTTTCACGTCACGACG
Plate: 51 row: N column: 8	Plate: 51 row: N column: 8	Plate: 51 row: N column: 8
Seq primer: ATTAGGTGACACTATG	Seq primer: ATTAGGTGACACTATG	Seq primer: ATTAGGTGACACTATG
Location/Qualifiers	Location/Qualifiers	Location/Qualifiers
1..555	1..555	1..555
/organism="Bos taurus"	/organism="Bos taurus"	/organism="Bos taurus"
/db_xref="taxon:9913"	/db_xref="taxon:9913"	/db_xref="taxon:9913"
/clone_lib="MARC 4BOV"	/clone_lib="MARC 4BOV"	/clone_lib="MARC 4BOV"
/tissue_type="pooled"	/tissue_type="pooled"	/tissue_type="pooled"
/lab_host="DH10B"	/lab_host="DH10B"	/lab_host="DH10B"
/note="vector: PCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from day 20 and day 40 embryos."	/note="vector: PCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from day 20 and day 40 embryos."	/note="vector: PCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from day 20 and day 40 embryos."
PAGE COUNT	PAGE COUNT	PAGE COUNT
130 a	130 a	130 a
157 c	157 c	157 c
133 g	133 g	133 g
135 t	135 t	135 t

ORIGIN	Query Match	24.4%	Score 186	DB 10	Length 555
	Best Local Similarity	67.2%	Pred. No. 4	4e-41	
	Matches 289	Conservative	0	Mismatches 120	Indels 21 Gaps 1
QY	12	CCACCAGACCCAGGCTGCCAGGCTGCATGTCAGTCTCTTGGCCATTTTCTGCATC	71		
DB	6	CAATCCAGACCCATGACTCCACGACGCTGCATCTGCTGACTTGGCCGCTGCTCATC	65		
QY	72	TCCAGGCTCTCTGTCACACGAGGCCCCAGTGCCCCCATGACTCTTAACTGATGCTG	131		
DB	66	CTTTGCTTCTCTGTCACATGCTGCTGCCCAATGTCCTCCACAGGACCTAGCTGTTGCTG	125		
QY	132	TGCCAGCCACACAGAGATGTGGGGACAAGTTCTTACAGCCCCCTGAGCAGCTGTGCTAT	191		
DB	126	TGCCATCATCAAGAGATGTGGGGAGCAAGTTCTTACAGCCCCCTGAGCAGCTGTGCTAT	185		
QY	192	GATGATGCGCTGTGCTGCTTGGCCAGGACCCAGACGCTGTGAACTGCACCTTCAGAGTC	251		
DB	186	GACGATGCCATGGTATCCCTTGGGACAGACCCGGAAGTGTGGAACTGCACCTTTACGGTC	245		
QY	252	TGCTTGGACAGAGTGGCGCCCTG-----GACCTTATGTTGAAG	290		
DB	246	TGCTTGGACAGTCTGCTGCCCTGTTGGTTCACAGGCCCATGATGATCTTCTGTGTGAAG	305		
QY	291	CTGATTAACACAGACTGCATCTAGCCCGACCTCGATGACAGGCTTGTGTCAGTGTG	350		
DB	306	GTTGAAGGTCAAAATGTTATCTCGGCTGTATCTCATAGATCGGGTGTGTGCGAGAGC	365		
QY	351	ACCTATGACATCATGAGGGAACGATGACTCTGTGATTCCTTCCCTGGTGGGCTCGA	410		
DB	366	AAACGAGGAGGACTCTGTGAATCTGCAACCTCAATATTTATTCATGAGGGGATGAT	425		
QY	411	GAAGAGGCT 420			
DB	426	GCAACGGGT 435			
RESULT 15					
LOCUS	BF911379	201 bp	mRNA	linear	EST 18-JAN-2001
DEFINITION	IL2-UT0073-081100-213-D02 UT0073		Homo sapiens cDNA, mRNA sequence.		
ACCESSION	BF911379				
VERSION	BF911379.1	GI:12302837			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheraia; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 201)				
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brumstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.				
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags				
JOURNAL MEDLINE COMMENT	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000) 2020263				
	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil				
	Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br				
	This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?fl=IL2&t2=IL2-UT0073-081100-213-D02&t3=2000-11-08&t4=1)				

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 23:01:37 ; Search time 48 Seconds
(without alignments)
3904.549 Million cell updates/sec

Title: US-09-092-297-5

Perfect score: 763
Sequence: 1 CTCGCTGACACACACACAG.....ATTAATTATGACTTTATATA 763

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 405 summaries

Database :

Issued Patents: NA: *
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PCRTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/lna/backfiles1.seq: *

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	33.2	4.4	6244	1 US-08-076-726-15	Sequence 15, Appl
C 2	33.2	4.4	6244	1 US-08-260-452-8	Sequence 8, Appl
C 3	33.2	4.4	6244	2 US-08-481-970-8	Sequence 8, Appl
C 4	33.2	4.4	6244	2 US-08-897-719-8	Sequence 8, Appl
C 5	33.2	4.4	6244	4 US-09-163-269-8	Sequence 8, Appl
C 6	32.2	4.2	9721	4 US-09-345-217-2	Sequence 2, Appl
C 7	32.2	4.2	152331	3 US-09-128-155-16	Sequence 16, Appl
C 8	32.2	4.2	176373	3 US-09-128-155-17	Sequence 17, Appl
C 9	31.8	4.2	963	4 US-08-438-313-126	Sequence 526, App
C 10	31.8	4.2	1474	1 US-08-465-980-1	Sequence 1, Appl
C 11	31.8	4.2	1474	2 US-09-053-303-1	Sequence 1, Appl
C 12	31.8	4.2	1474	5 PCT-US95-07093-1	Sequence 1, Appl
C 13	31.2	4.1	798	4 US-09-199-637A-30	Sequence 30, Appl
C 14	31.2	4.1	1407	4 US-09-199-637A-28	Sequence 28, Appl
C 15	31.2	4.1	3141	4 US-09-199-637A-12	Sequence 12, Appl
C 16	31.2	4.1	42235	4 US-09-199-637A-1	Sequence 1, Appl
C 17	31	4.1	2565	1 US-08-597-495B-29	Sequence 29, Appl
C 18	31	4.1	931	1 US-09-068-051A-31	Sequence 31, Appl
C 19	30.8	4.0	1893	3 US-08-203-905B-4	Sequence 4, Appl
C 20	30.8	4.0	2038	3 US-08-878-474-10	Sequence 10, Appl
C 21	30.8	4.0	3900	2 US-08-276-531-45	Sequence 45, Appl
C 22	30.8	4.0	3900	2 US-08-356-060A-42	Sequence 42, Appl
C 23	30.8	4.0	3900	4 US-08-460-900C-42	Sequence 42, Appl
C 24	30.8	4.0	3900	4 US-08-674-509B-42	Sequence 42, Appl
C 25	30.8	4.0	3900	4 US-08-954-698-42	Sequence 42, Appl
C 26	30.8	4.0	4434	2 US-08-540-405-5	Sequence 5, Appl
C 27	30.8	4.0	4434	3 US-08-656-055-5	Sequence 5, Appl

C 28	30.8	4.0	4434	4 US-08-954-668-5	Sequence 5, Appl
C 29	30.8	4.0	4434	5 PCT-US95-13233-5	Sequence 5, Appl
C 30	30.4	4.0	2363	4 US-09-721-383-1	Sequence 1, Appl
C 31	30.2	4.0	7218	1 US-08-232-463-14	Sequence 14, Appl
C 32	30.2	4.0	9559	3 US-08-014-416-2	Sequence 2, Appl
C 33	30.2	4.0	9559	3 US-09-014-416-6	Sequence 6, Appl
C 34	29.8	3.9	508	3 US-09-058-389A-21	Sequence 21, Appl
C 35	29.8	3.9	2396	3 US-09-058-389A-10	Sequence 10, Appl
C 36	29.8	3.9	10367	1 US-08-110-300A-9	Sequence 9, Appl
C 37	29.8	3.9	10367	2 US-08-886-642-9	Sequence 9, Appl
C 38	29.8	3.9	10367	2 PCT-US93-08041-9	Sequence 9, Appl
C 39	29.6	3.9	289	4 US-09-007-005-17	Sequence 17, Appl
C 40	29.6	3.9	289	4 US-09-244-796-17	Sequence 17, Appl
C 41	29.6	3.9	2712	2 US-08-410-784A-1	Sequence 1, Appl
C 42	29.4	3.9	1055	2 US-08-828-242-2	Sequence 2, Appl
C 43	29.4	3.9	1055	4 US-09-206-499-2	Sequence 2, Appl
C 44	29.4	3.9	1083	1 US-08-592-126-75	Sequence 75, Appl
C 45	29.4	3.9	1110	4 US-09-342-084-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-076-726-15/c
Sequence 15, Application US/08076726
Patent No. 5464758
GENERAL INFORMATION:
APPLICANT: Gossen, Manfred
APPLICANT: Bujard, Hermann
TITLE OF INVENTION: Tight Control of Gene Expression in
TITLE OF INVENTION: Eucaryotic Cells by Tetraacycline-responsive Promoters
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: District of Columbia
COUNTRY: United States of America
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/076,726
FILING DATE: 14-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942, 2490001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
TELEX: 248636 SRK
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 6244 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-076-726-15
Query Match 4.4%: Score 33.2; DB 1; Length 6244;
Best Local Similarity 48.9%: Pred. No. 5.8;
Matches 89; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
352 GCTAATGACATCAGGGAAGATGACTCTGATTCCTGCTGAGGCGCTGAG 411
DB 2733 GTTCATGCTCATTAACAGGTGATTAGAGGGGAATTAAGTATCTGACTTGAG 2674

OY	412	AAAGAGCTGGTGTACCTAGATCTGGAATCGTGAGTCGCTTTGGGCGCCACAAGAAA	471
Db	2673	AAAAAGTATTCTGTTGGCTTTCATTTTGGAATGCCCACTGGCTGTGGAGAAGCAACAGCAT	261
OY	472	CACACACTCAACTGCCACATTCAATTCTGTGTAAGCTGTCTGAGGCCAACCTGGCGCTGCC	531
Db	2613	CGAGTGCCTCAGTCAGTCCTGACTTTATTTGAACTTTTAAACTTTTGGCCTCCAGAAGACCA	255
OY	532	TG 533	
Db	2553	TG 2552	

```

RESULT 2
US-08-260-452-8/c
: Sequence 8, Application US/08260452
: Patent No. 5650298
: GENERAL INFORMATION:
: APPLICANT: Gossen, Manfred
: APPLICANT: Bujard, Hermann
: APPLICANT: Salfeld, Jochem
: APPLICANT: Voss, Jeffrey
: TITLE OF INVENTION: Tight Control of Gene Expression in Eucaryotic
: TITLE OF INVENTION: Cells by Tetracycline-responsive Promoters
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lohive & Cockfield
: STREET: 60 State Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109-1875
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/260,452
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/076,327
: ATTORNEY/AGENT INFORMATION:
: NAME: Giulio A. Decontt, Jr.
: REGISTRATION NUMBER: 31,503
: REFERENCE/DOCKET NUMBER: BBI-013CP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 227-5941
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6244 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: circular
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: Human cytomegalovirus
: STRAIN: Towne (hcmv)
: IMMEDIATE SOURCE:
: CLONE: PUHD BGR3
: US-08-260-452-8

Query Match 4.4% Score 33.2; DB 1; Length 6244;
Best Local Similarity 48.9%; Pred. No. 5.8;
Matches 89; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 352 GCTAATGACATCATGAGGGAACGATGACTCCTGATTCCTCCTGAGGAGCCTGGAG 411
Db 2733 GTTCATAGCTCATTAACAGCGTGATTAAGAGGGGAGATTAACTGATCTCTTGACTTGGAG 2674
Y 412 AAGAGGCTGGTGTACTCTGATGATCTGGAGTCTGAGTGGCTGTTTGGGGCCAGAGAA 471

```

[illegible]

```

US-08-481-970-8/c
Sequence 8, Application US/08481970
Patent No. 5859310
GENERAL INFORMATION:
APPLICANT: Gossen, Manfred
APPLICANT: Bujard, Hermann
APPLICANT: Safeld, Udochen
APPLICANT: Voss, Jeffrey
TITLE OF INVENTION: Aitmal Transgenic for a tetracycline-Controlled
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,970
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/260,452
FILING DATE: 14-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076,327
FILING DATE: 14-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. Decontil, Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-013CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 6244 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Human cytomegalovirus
STRAIN: Towne (hcmv)
IMMEDIATE SOURCE:
CLONE: PUHD BGR3
US-08-481-970-8

Query Match 4.4%; Score 33.2; DB 2; Length 6244;
Best Local Similarity 48.9%; Pred. No. 5.8;
Matches 89; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

352 GCTAATGACATAGAGGGAACGATGACTCTCGATTCTCCTTCTGGTGGCCCTGGAG 411
b 2733 GTTCAAGCCTCATTAACAGGTTAGAGGGGCAATTAACTGATCTCTTGACTTGGAG 2674

```


QY 133 GCCAGCCACACAGAGATGTGGGACACTTCTACGACCCCTGC 177
DB 138170 GCCTGAACACAGATGCTGGAGCGATTAATGTGCATCCCTCC 138126

RESULT 9

US-09-439-313-526/c
Sequence 526, Application US/09439313

Patent No. 6329505

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, David C.

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan Louise

APPLICANT: Jiang Yuqul

APPLICANT: Reed, Steven G.

APPLICANT: Kalos, Michael

APPLICANT: Fanger, Gary

APPLICANT: Retter, Mark

APPLICANT: Solk, John

APPLICANT: Day, Craig

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND

FILE REFERENCE: 210121.427C9

CURRENT APPLICATION NUMBER: US/09/439,313

CURRENT FILING DATE: 1999-11-12

NUMBER OF SEQ ID NOS: 575

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO: 526

LENGTH: 963

TYPE: DNA

ORGANISM: Homo sapiens

US-09-439-313-526

Query Match 4.2%; Score 31.8; DB 4; Length 963;
Best Local Similarity 49.1%; Pred. No. 7.2;

Matches 84; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 93 GGAGCCCGAGTGGCCCGCATGACTCTTACCTGATCTGTGCGACGACACAGAGATGT 152
DB 510 GGAGTGGCGAAGAGCGCCGCTGTGATCAGCAGAGCAGTGGGAAAAAAGAGGATCC 451
QY 153 GGGGACAGTCTACGACCCCTGACGACCTGTGCTATGATGATGATGATGATGATG 212
DB 450 GCGGACACAGCAGCAGATGCAATCTGGGCTGTACTGTATTGTTGAGCAGCTGACATG 391
QY 213 GCCAGACCCAGACAGTGTGGAACCTGACCTTACAGATCTGCTTTGAGCAG 263
DB 390 GCGGAGTGGGTGGCAGATGGCCACATACGGTCAAGGCCATGGCCAGCAG 340

RESULT 10

US-08-465-980-1/c

Sequence 1, Application US/08465980

Patent No. 5756309

GENERAL INFORMATION:

APPLICANT: Soppet, Daniel R.

APPLICANT: Li, Yi

APPLICANT: Rosen, Craig A.

APPLICANT: Ruben, Steven M.

TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAT70

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,

STREET: 6 Becker Farm Road

CITY: Roseland

STATE: New Jersey

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/465,980

FILING DATE: 06-JUN-1995

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Ferrari, Gregory D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-446

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1474 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 274..1233

US-08-465-980-1

Query Match 4.2%; Score 31.8; DB 1; Length 1474;
Best Local Similarity 49.1%; Pred. No. 8.4;

Matches 84; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 93 GGAGCCCGAGTGGCCCGCATGACTCTTACCTGATCTGTGCGACGACACAGAGATGT 152
DB 783 GGAGTGGCGAAGAGCGCCGCTGTGATCAGCAGAGCAGTGGGAAAAAAGAGGATCC 724
QY 153 GGGGACAGTCTACGACCCCTGACGACCTGTGCTATGATGATGATGATGATGATG 212
DB 723 GCGGACACAGCAGCAGATGCAATCTGGGCTGTACTGTATTGTTGAGCAGCTGACATG 664
QY 213 GCCAGACCCAGACAGTGTGGAACCTGACCTTACAGATCTGCTTTGAGCAG 263
DB 663 GCGGAGTGGGTGGCAGATGGCCACATACGGTCAAGGCCATGGCCAGCAG 613

RESULT 11

US-09-053-303-1/c

Sequence 1, Application US/09053303

Patent No. 5948890

GENERAL INFORMATION:

APPLICANT: Soppet, Daniel R.

APPLICANT: Li, Yi

APPLICANT: Rosen, Craig A.

APPLICANT: Ruben, Steven M.

TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAT70

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,

STREET: 6 Becker Farm Road

CITY: Roseland

STATE: New Jersey

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/053,303

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/465,980

FILED DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-446
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1474 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 274..1233
US-09-053-303-1

Query Match
Best Local Similarity 4.2%; Score 31.8; DB 2; Length 1474;
Matches 84; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

OY 93 GGAGCCCCAGTGGCCCCCATGACTCTTACCTGATGCTGTGCCAGCCACACAAGAGTGT 152
DB 783 GGAGTGGCAGAAAGCCAGCCGCTTGATCAGCAGAGCAGTGGGAAAAAAGAGGATCC 724
OY 153 GGGGACAAGTTCTACGACCCCTCGACGACCTGTTCTATGATGATGACCGTGTGCCCTTG 212
DB 723 GCGGACACAGCCAGCAGATGCGCAATCTGGCTGTACTGATTTGTGAGCAGCTGCAGCATG 664
OY 213 GCCAGGACCCAGACCTGTGGAACCTGCACCTTCAGAGTCTCTTTGAGCAG 263
DB 663 GCGAGTGGTGGCAGATGCGCCACATACGTCATAAGGCCATGGCCAGCAG 613

RESULT 12

PCT-US95-07093-1/c

Sequence 1, Application PC/TUS9507093

GENERAL INFORMATION:

APPLICANT: Soppet, Daniel R.

APPLICANT: Li, Yi

APPLICANT: Rosen, Craig A.

APPLICANT: Ruden, Steven M.

TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAL70

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARIELLA, BYRNE, BAIN, GILFILLAN, CECCHI,

STREET: 6 Becker Farm Road

CITY: Roseland

STATE: New Jersey

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/07093

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/465,980

FILING DATE: 06-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Ferraro, Gregory D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-446

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1474 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 274..1233
PCT-US95-07093-1

Query Match
Best Local Similarity 4.2%; Score 31.8; DB 5; Length 1474;
Matches 84; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

OY 93 GGAGCCCCAGTGGCCCCCATGACTCTTACCTGATGCTGTGCCAGCCACACAAGAGTGT 152
DB 783 GGAGTGGCAGAAAGCCAGCCGCTTGATCAGCAGAGCAGTGGGAAAAAAGAGGATCC 724
OY 153 GGGGACAAGTTCTACGACCCCTCGACGACCTGTTCTATGATGATGACCGTGTGCCCTTG 212
DB 723 GCGGACACAGCCAGCAGATGCGCAATCTGGCTGTACTGATTTGTGAGCAGCTGCAGCATG 664
OY 213 GCCAGGACCCAGACCTGTGGAACCTGCACCTTCAGAGTCTCTTTGAGCAG 263
DB 663 GCGAGTGGTGGCAGATGCGCCACATACGTCATAAGGCCATGGCCAGCAG 613

RESULT 13

US-09-199-637A-30

Sequence 30, Application US/09199637A

Patent No. 635411

GENERAL INFORMATION:

APPLICANT: Ausubel, Frederick

APPLICANT: Goodman, Howard M.

APPLICANT: Rahme, Laurence G.

APPLICANT: Mahajan-Miklos, Shalina

APPLICANT: Cao, Hui

APPLICANT: Drenkard, Eliana

APPLICANT: Tsongalis, John

TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID

FILE REFERENCE: 00786/361002

CURRENT APPLICATION NUMBER: US/09/199,637A

CURRENT FILING DATE: 1998-11-25

PRIOR APPLICATION NUMBER: 60/066,517

PRIOR FILING DATE: 1997-11-25

NUMBER OF SEQ ID NOS: 437

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 30

LENGTH: 798

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-199-637A-30

Query Match
Best Local Similarity 4.1%; Score 31.2; DB 4; Length 798;
Matches 66; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

OY 89 ACAGGAGCCCCAGTGGCCCCCATGACTCTTACCTGATGCTGTGCCAGCCACACAAGAG 148
DB 301 ATACTGCTGCTGCTGTGTTCCCGAGTCTTACTTGCCCTTGCCATGACCGAAGAGA 360
OY 149 ATGTGGGACAAGTTCTACGACCCCTCGACGACCTGTTCTATGATGATGACCGTGTGCC 208
DB 361 AGAAGACACAGCCAGCAGATGCGCAATCTGGCTGTACTGATTTGTGAGCAGCTGCAGCATG 420
OY 209 CTTG 212
DB 421 CTTG 424

RESULT 14

US-09-199-637A-28/c
Sequence 28, Application US/09199637A
Patent No. 6355411
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick
APPLICANT: Goodman, Howard M.
APPLICANT: Rahme, Laurence G.
APPLICANT: Mahajan-Miklos, Shalina
APPLICANT: Tan, Man-Wah
APPLICANT: Cao, Hui
APPLICANT: Drenkard, Eliana
APPLICANT: Tsongalis, John
TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
FILE REFERENCE: 00786/361002
CURRENT APPLICATION NUMBER: US/09/199,637A
PRIOR FILING DATE: 1998-11-25
NUMBER OF SEQ ID NOS: 437
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28
LENGTH: 1407
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-28

Query Match 4.1%; Score 31.2; DB 4; Length 1407;

Best Local Similarity 53.2%; Pred. No. 12;
Matches 66; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 89 ACACGAGCCCCAGTGGCCCCCATGACTCTTACCTGATGCTGCGCAGCCACAAG 148
DB 448 ATACCTCGTCGCTGTGTGTCCTCCCGAGTCTTACTGCGCTGGCCATGACGAAACGA 389
QY 149 ATGTGGGACAACTTCTACGACCCCTGCGAGCACTGTCTATGATGATGCCCTGCGC 208
DB 388 AGAAAGAACCGAGCGCTACACATCATGCAAGCAGCAGCGCTGCGAGAGCTCGAGCGGC 329
QY 209 CTTG 212
DB 328 CTTG 325

RESULT 15

US-09-199-637A-12
Sequence 12, Application US/09199637A
Patent No. 6355411
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick
APPLICANT: Goodman, Howard M.
APPLICANT: Rahme, Laurence G.
APPLICANT: Mahajan-Miklos, Shalina
APPLICANT: Tan, Man-Wah
APPLICANT: Cao, Hui
APPLICANT: Drenkard, Eliana
APPLICANT: Tsongalis, John
TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
FILE REFERENCE: 00786/361002
CURRENT APPLICATION NUMBER: US/09/199,637A
PRIOR FILING DATE: 1998-11-25
NUMBER OF SEQ ID NOS: 437
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 3141
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa

US-09-199-637A-12

Query Match 4.1%; Score 31.2; DB 4; Length 3141;
Best Local Similarity 53.2%; Pred. No. 17;
Matches 66; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 89 ACACGAGCCCCAGTGGCCCCCATGACTCTTACCTGATGCTGCGCAGCCACAAG 148
DB 2925 ATACCTCGTCGCTGTGTGTCCTCCCGAGTCTTACTGCGCTGGCCATGACGAAACGA 2984
QY 149 ATGTGGGACAACTTCTACGACCCCTGCGAGCACTGTCTATGATGATGCCCTGCGC 208
DB 2985 AGAAAGAACCGAGCGCTACACATCATGCAAGCAGCAGCGCTGCGAGAGCTCGAGCGGC 3044
QY 209 CTTG 212
DB 3045 CTTG 3048

Search completed: November 2, 2002, 23:18:35
Job time: 257 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OW protein - nucleic search, using frame_plus.p2n model

Run on: November 3, 2002, 03:02:24 : Search time 1822 Seconds

(Without alignments)
1343.801 Million cell updates/sec

Title: US-09-092-297-17

Perfect score: 668
Sequence: 1 PLOPPRMAPRGCTIVAFAL.....INQCDARSRTDRLCRSVS 117

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ .p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US09092297/rnat.29102002.093923.21750/app.query.fasta.1.263
-DB=GenEmbl -GFM=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEC=0 -LOOPEXT=0
-UNITS=bits -STAR=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09092297.GCGN.1.1.1182.@rnat.29102002.093923.21750 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_htg:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*
15: em_ba:.*
16: em_fun:.*
17: em_hum:.*
18: em_in:.*
19: em_mu:.*
20: em_cm:.*
21: em_or:.*
22: em_ov:.*
23: em_ph:.*
24: em_pat:.*
25: em_pl:.*
26: em_ro:.*
27: em_sts:.*
28: em_un:.*

29: em_vl:.*
30: em_htg_hum:.*
31: em_htg_inv:.*
32: em_htg_other:.*
33: em_htgo_inv:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	668	100.0	762	6	AX014140
2	668	100.0	771	6	AX164131
3	531.5	79.6	107573	9	AC006262
4	289	43.3	146180	9	AC007193
5	237	35.5	662	6	AX055440
6	237	35.5	662	6	AX055696
7	236	35.3	378	6	AX118905
8	197.5	29.6	166500	9	AC007785
9	194	29.0	166500	9	AC007785
10	94.5	14.1	47407	2	AC094931
11	94.5	14.1	115627	2	AC094152
12	91	13.6	158325	2	AC087564
13	89.5	13.4	10197	1	AE005499
14	89.5	13.4	270365	1	AP002562
15	89	13.3	206850	10	AC073761
16	89	13.3	222610	10	AC026385
17	89	13.3	232190	2	AC074167
18	88.5	13.2	163259	2	AL645992
19	88.5	13.2	170273	9	CNS057CD
20	88.5	13.2	325791	6	AX234657
21	88	13.2	1907	5	FRCB1B
22	87.5	13.1	1952	10	MMU76007
23	87.5	13.1	161044	9	AC011746
24	87.5	13.1	193046	2	AC023348
25	87.5	13.1	210869	2	AC109608
26	87	13.0	14743	10	AF270884
27	87	13.0	129389	9	HS326114
28	86.5	12.9	106578	9	AL137249
29	86.5	12.9	117927	9	AC013416
30	86.5	12.9	157742	2	AP000385
31	86.5	12.9	184192	2	AP001154
32	86.5	12.9	189717	9	AC090952
33	86	12.9	16556	9	AF176832
34	86	12.9	35530	3	CEM05B2
35	86	12.9	45346	2	AC095880
36	86	12.9	161866	2	AC103301
37	86	12.9	163388	2	AC096078
38	86	12.9	185195	2	AC074001
39	86	12.9	199957	2	AL589735
40	86	12.9	214603	2	AC098941
41	86	12.9	262426	2	AL645533
42	85.5	12.8	77543	2	AC095765
43	85.5	12.8	115199	9	AC003695
44	85.5	12.8	119770	2	AC008697
45	85.5	12.8	148334	2	AC087106

ALIGNMENTS

RESULT 1
LOCUS AX014140
DEFINITION Sequence 3 from Patent WO9554447.
ACCESSION AX014140
VERSION AX014140.1 GI:10040587
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

DNA linear PAT 07-SEP-2000

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
Schmitt, A., Specht, T., Dahl, E., Hinmann, B., Rosenthal, A. and
Pillarsky, C.
TITLE
Human nucleic acid sequences of bladder tumour tissue
JOURNAL
Patent: WO 954447-A 3 28-OCT-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINNMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILLARSKY CHRISTIAN (DE)
FEATURES
location/Qualifiers
source
1. 762
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT
162 a 232 c 194 g 174 t
ORIGIN

Alignment Scores:
Pred. No.: 6,14e-61 Length: 762
Score: 668.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 6

US-09-092-297-17 (1-117) x AX014140 (1-762)

QY 1 ProleuglnPropoArgAlaMetAlaProArgGlyCysIleValAlaValAlaPheAlaIle 20
Db 3 CCACTGCACACCCACAGCCATGGCTCCCGAGGCTGCATCGTACGTCTTGGCCATT 62
QY 21 PheArgYAlcSerArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 63 TTTCAGATCTCCAGCTCTCTGCTCAGACGAGAGCCAGGCCCCCATGACCTCTTAC 122
QY 41 LeuMetLeuGlyGlnProHisLysArgCysGlyAspLysPheTyrAspProLeuGlnHis 60
Db 123 CTGATGCTGTGCGCCACACACAGATGTGGGACAAAGTTCTACAGACCCCTGCAGCAC 182
QY 61 CysCysTyrAspAspAlaValAlaValProLeuAlaArgThrGlnThrCysGlyAsnCysThr 80
Db 183 TGTTCATATGATGATGCGGTGCTCCCTTGGCCAGACCCAGACGTGTGGAATACTGCACC 242
QY 81 PheArgYAlcSerArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
Db 243 TTTCAGATCTCTGCTTGGAGCAGTGTGCTCCCTGCAGCTTCATGAGTGAAGATGAACAC 302
QY 101 AsnCysAspSerAlaArgThrSerAspAspArgLeuGlyAsnGlySerValSer 117
Db 303 AACTGCAGCTCAGCCCGGACCTCGATGACAGACGCTTGTGTCAGATGTGCACG 353

RESULT 2
AX164131 771 bp DNA 1linear PRI 22-JUN-2001
LOCUS
DEFINITION
Sequence 3 from Patent WO0140465.
ACCESSION
AX164131
VERSION
AX164131.1 GI:14545080
KEYWORDS
human.
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
Fong, S., Goddard, A., Godowski, P., J., Grimaldi, C., J., Gurney, A. L.,
Hillan, K. J., Tuma, D., Matanabe, C. K., Wood, W. I. and Zhang, Z.
TITLE
Compositions and methods for the treatment of immune related
diseases
JOURNAL
Patent: WO 0140465-A 3 07-JUN-2001;
Genentech, Inc. (US)
FEATURES
location/Qualifiers
source
1. 771
/organism="Homo sapiens"
/db_xref="taxon:9606"

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
Schmitt, A., Specht, T., Dahl, E., Hinmann, B., Rosenthal, A. and
Pillarsky, C.
TITLE
Human nucleic acid sequences of bladder tumour tissue
JOURNAL
Patent: WO 954447-A 3 28-OCT-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINNMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILLARSKY CHRISTIAN (DE)
FEATURES
location/Qualifiers
source
1. 762
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT
169 a 231 c 195 g 176 t
ORIGIN

Alignment Scores:
Pred. No.: 6,21e-61 Length: 771
Score: 668.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 6

US-09-092-297-17 (1-117) x AX164131 (1-771)

QY 1 ProleuglnPropoArgAlaMetAlaProArgGlyCysIleValAlaValAlaPheAlaIle 20
Db 3 CCACTGCACACCCACAGCCATGGCTCCCGAGGCTGCATCGTACGTCTTGGCCATT 62
QY 21 PheArgYAlcSerArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 63 TTTCAGATCTCCAGCTCTCTGCTCAGACGAGAGCCAGGCCCCCATGACCTCTTAC 122
QY 41 LeuMetLeuGlyGlnProHisLysArgCysGlyAspLysPheTyrAspProLeuGlnHis 60
Db 123 CTGATGCTGTGCGCCACACACAGATGTGGGACAAAGTTCTACAGACCCCTGCAGCAC 182
QY 61 CysCysTyrAspAspAlaValAlaValProLeuAlaArgThrGlnThrCysGlyAsnCysThr 80
Db 183 TGTTCATATGATGATGCGGTGCTCCCTTGGCCAGACCCAGACGTGTGGAATACTGCACC 242
QY 81 PheArgYAlcSerArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
Db 243 TTTCAGATCTCTGCTTGGAGCAGTGTGCTCCCTGCAGCTTCATGAGTGAAGATGAACAC 302
QY 101 AsnCysAspSerAlaArgThrSerAspAspArgLeuGlyAsnGlySerValSer 117
Db 303 AACTGCAGCTCAGCCCGGACCTCGATGACAGACGCTTGTGTCAGATGTGCACG 353

RESULT 3
AC006262 107573 bp DNA 1linear PRI 21-SEP-2001
LOCUS
DEFINITION
Homo sapiens chromosome 19, BAC BC264576 (CTC-29/N14), complete
sequence.
ACCESSION
AC006262
VERSION
AC006262.2 GI:15718550
KEYWORDS
HTG.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
Lamerdin, J. E., McCready, P. M., Skowronski, E., Visswanathan, V.,
Burkhardt-Schultz, K., Gordon, L., Dias, J., Ramirez, M., Stilaugen, S.,
Phan, H., Velasco, N., Do, L., Regala, W., Terry, A., Ganes, J.,
Dangnan, L., Erler, A., Christensen, M., Georgescu, A., Aylla, J.,
Liu, S., Altix, C., Andreise, T., Trankheim, M., Amico-Feller, G.,
Coefield, J., Duarte, S., Lucas, S., Bruce, R., Thomas, P., Quan, G.,
Krommiller, B., Ariellano, A., Sanders, C., Ow, D., Nolan, M., Trong, S.,
Kobayashi, A., Olsen, A. S. and Carraro, A. V.
TITLE
Sequence analysis of a 1.9 Mb region in 19q13.2 between APOE and
D19S412
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 107573)
AUTHORS
Lamerdin, J. E.
TITLE
Direct Submission
JOURNAL
Submitted (30-DEC-1998) Joint Genome Institute, Lawrence Livermore
National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
3 (bases 1 to 107573)
AUTHORS
Lamerdin, J. E.
TITLE
Direct Submission
JOURNAL
Submitted (21-SEP-2001) Joint Genome Institute, Lawrence Livermore
National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
On Sep 21, 2001 this sequence version replaced gi:4079612.
Map and sequence oriented from centromere to q telomere. BC264576

```

overlaps BC282485 (CTC-344H19, AC007785) to the left from bases 1 to 20,296 of this accession, and overlaps BC82621 (CTB-139A18, AC007193) to the right from bases 106,317 to 107,573 of this accession. This sequence does not represent the entire insert of BC24576. Additional chromosome 19 map and sequence information may be obtained at: <http://www-bio.11n1.gov/dbp/genome/genome.html>.

FEATURES

Source

```

1..107573
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2 between APOE and D19S412"
/clone="CTC-297N14"
/cell_line="978SK"
/cell_type="fibroblast"
/note="LLNL clone name: BC264576 BAC library obtained from Research Genetics."
354..567
/note="DSS similarity to AA456254 zx99d06.r1 Soares NHHMPUS1 Homo sapiens CDNA clone 811883 5'; (3..215); 100% identity."
repeat_region
/rpt_family="GA)n"
/complement(669..800)
repeat_region
/rpt_family="(GA)n"
/complement(829..953)
repeat_region
/rpt_family="(GGCA)n"
/complement(1035..2017)
repeat_region
/rpt_family="L1MB7"
/complement(2014..2202)
repeat_region
/rpt_family="L1MA7"
/complement(2206..2495)
repeat_region
/rpt_family="AluJo"
/complement(2497..2728)
repeat_region
/rpt_family="L1MA7"
/complement(2730..3133)
repeat_region
/rpt_family="L1MB7"
3158..3447
/rpt_family="AluJb"
/complement(3448..3803)
repeat_region
/rpt_family="L1MB7"
/complement(3805..4108)
repeat_region
/rpt_family="AluJo"
/complement(4110..4801)
repeat_region
/rpt_family="L1MB7"
/complement(4826..5069)
repeat_region
/rpt_family="L1M1"
/complement(5338..5369)
repeat_region
/rpt_family="(TA)n"
/complement(5370..5708)
repeat_region
/rpt_family="L1PA8"
/complement(5974..6277)
repeat_region
/rpt_family="AluY"
/complement(7059..7080)
repeat_region
/rpt_family="AT-rich"
/complement(7150..7485)
repeat_region
/rpt_family="L1"
/complement(7527..7646)
repeat_region
/rpt_family="FLAM_C"
/complement(7661..7916)
repeat_region
/rpt_family="L1"
7938..9725
/rpt_family="L1"
/complement(9736..10038)
repeat_region
/rpt_family="AluSg"
/complement(10866..10964)
/note="Predicted exon, program: graal2exons_human_1.3, frame: 1, quality: excellent, score: 75.000"
11326..11517
/rpt_family="AluJo"
11945..11998
/note="DSS similarity to AA456254 zx99d06.r1 Soares NHHMPUS1 Homo sapiens CDNA clone 811883 5'; (216..265); 100% identity."
12144..12411
misc_feature

```

```

/note="DSS similarity to overlapping ESTs: (12144..12398) AA456254 zx99d06.r1 Soares NHHMPUS1 Homo sapiens CDNA clone 811883 5'; (270..523); 99% identity." (12182..12411) A1081371 ox76g10.x1 Soares NHHMPUS1 Homo sapiens CDNA clone IMAGE:1662306 3'; (454..225) 100% identity." (12173..12411) AA454642 zx99d06.r1 Soares NHHMPUS1 Homo sapiens CDNA clone 811883 3'; (455..217); 100% identity." (12144..12411) AA760247 vv74e06.r1 StrataGene mouse skin (#937313) Mus musculus CDNA clone 1228162 5'; (119..380); 69% identity."
complement(12536..12614)
/note="Predicted exon, program: graal2exons_human_1.3, frame: 2, quality: excellent, score: 90.000"
12600..12818
/note="DSS similarity to overlapping ESTs: (12603..12818) AA454642 zx99d06.r1 Soares NHHMPUS1 Homo sapiens CDNA clone 811883 3'; (216..1); 99% identity." (12600..12823) A1081371 ox76g10.x1 Soares NHHMPUS1 Homo sapiens CDNA clone IMAGE:1662306 3'; (224..1); 98% identity." (12600..12656) AA760247 vv74e06.r1 StrataGene mouse skin (#937313) Mus musculus CDNA clone 1228162 5'; (381..432); 76% identity."
complement(12967..16062)
/rpt_family="L1P5"
/complement(14029..14445)
/note="Predicted exon, program: graal2exons_human_1.3, frame: 0, quality: good, score: 58.000"
complement(17695..18620)
/rpt_family="MER11B"
/complement(18619..18767)
repeat_region
/rpt_family="MER11A"
19151..19519
/rpt_family="MLT1B"
/complement(19592..19754)
repeat_region
/rpt_family="AluJo"
/complement(19767..19870)
repeat_region
/rpt_family="MER65A"
/complement(20427..20481)
repeat_region
/rpt_family="(CA)n"
20546..20751
/rpt_family="AluJo"
21128..21397
/rpt_family="AluSx"
21768..22798
/rpt_family="MER11A"
/complement(23090..23194)
/note="Predicted exon, program: graal2exons_human_1.3, frame: 2, quality: good, score: 50.000"
complement(24023..24141)
/rpt_family="FLAM_A"
/complement(24258..24429)
repeat_region
/rpt_family="AluJb"
/complement(24910..25309)
repeat_region
/rpt_family="MSTB"
25329..25464
/rpt_family="HERVL"
25732..25881
/rpt_family="MER45"
/complement(26586..26613)
repeat_region
/rpt_family="AT-rich"
/complement(26677..26941)
repeat_region
/rpt_family="MIR"
/complement(27078..27369)
repeat_region
/rpt_family="AluSx"
complement(27374..27682)
/rpt_family="AluSg"
27743..28177
/rpt_family="L1M1/2"
28646..28929
/rpt_family="AluSx"
28967..29070
/rpt_family="(TA)n"
29085..29241
misc_feature

```

Alignment Scores:	
Pred. No.:	2,3e-44
Score:	531.50
Percent Similarity:	44.31%
Best Local Similarity:	44.31%
Query Match:	79.57%
DB:	9
Length:	107573
Matches:	113
Conservative:	0
Mismatches:	1
Indels:	142
Gaps:	2

QY 1 PROLEUGLPROPROARGALAMETALAProARGGLYCYSILEVAL----- 15

— — — — —

DB 81284 CCAC TGC AAC CACC AGAGCC ATGGCT CCCG AGGCT GCATCGT -AGGT AAGAG GACAG 81

QY 15 ----- 15

Db 81343 GACCCCATTCACCTGAGCCCATCTCCATCTCAGCCACCACCAGGCTCTCTCCCTAC 81

0V 15 ----- 15

Db 81403 CCCAACTACACCTCATCCCTGTTCCCATGCCAGCCAAATCTCCAGGCGTGAATTGCTC 81

0x 15 ----- 1e

13

Db 81463 ACAGCAATCCCATCATAGTGCCCAATCAGCCCATCCACAGCCCTGTTCACTCC 81

15 -----

15

Db 81523 TGAGCTGCTCAAAGCCATACCGAGCCCTGGTTCACACCAGTCTTCTTAATGTCAGTG 81

Curriculum

AlaVal 17

D_b 81583 CCACTCATATCTGTGGCTGTCACACTCACCCCATCTTCCCCTTTCCCCTTTCCACAGCTGTC 81
 |||||

18 Phobias, Anxiety Disorders, and the

16 pncysllesatrgleuleucysserhsglyala----- 32

Db 81643 TTGCCATTCTGCATCTCCAGGCTCCTCTGCTCACACGGAGCCCAAGTGAGCCCAAG 81

3

32 ----- 32

Db 81703 AGTGTTCGGAAGCTGGAGGAGGGGTACACCTTCCAGGGGACCCAGGCTGGGACCCACC 81

1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities related to the business. It emphasizes the need for transparency and accountability in financial reporting.

33 -----proValAlaPromethThrProty 40

Db 81763 TGGCACTACCTCTCGATCTCACCAGCTCTGTCTCTCCAGTGGCCCCATGACTCCTTA 81

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

40 IleuMetLeuGlyGlnProHisLysArgGlyAspLysPheTyrAspProLeuGlnH₁ 60

Db 81823 CCTGATGCTGTGCCAGCCACACAGAGATGTGGGACAAGTCTACGACCCCTGCAGCA 81

A

60 scyscystyraspaalavalvalproleualargthrgrlntthrcysglyasnrcysth 80

.....

repeat_region

```

repeat_region /rpt_family="AluJo"
3355..3655
/rpt_family="AluSg"
repeat_region /rpt_family="AluSg"
3673..3939
/rpt_family="L1"
complement(4005..4264)
/rpt_family="AluSx"
repeat_region complement(4269..4568)
/rpt_family="AluY"
5064..5122
/Note="Predicted exon, program: g1all2exons_human_1.3,
frame: 2, quality: excellent, score: 81.000"
repeat_region /rpt_family="AluY"
5626..5886
/Note="Predicted exon, program: g1all2exons_human_1.3,
frame: 0, quality: good, score: 52.000"
misc_feature
5626..5886
repeat_region /rpt_family="Alu"
7398..7444
/rpt_family="Alu"
7480..7595
/rpt_family="(TA)n"
repeat_region /rpt_family="AluDb"
7596..7812
/rpt_family="AluDb"
7900..8191
/rpt_family="AluSx"
complement(8575..8597)
/rpt_family="AT-rich"
complement(8598..8878)
/rpt_family="AluSx"
9477..9766
/rpt_family="AluSx"
9997..10258
/rpt_family="AluSg"
10273..10347
/rpt_family="FLAM"
10691..10818
/rpt_family="AluDb"
complement(10999..11283)
/rpt_family="AluJo"
complement(11302..11517)
/rpt_family="L1MC/D"
complement(11536..11661)
/rpt_family="FLAM_C"
repeat_region complement(11667..11798)
/rpt_family="L1MC/D"
complement(11833..12138)
/rpt_family="AluSg"
complement(12158..12478)
/rpt_family="L1MC/D"
12479..12642
/rpt_family="FRAM"
complement(12655..12795)
/rpt_family="L1MC/D"
complement(12806..13106)
/rpt_family="AluSx"
complement(13356..13481)
/rpt_family="(GAA)n"
complement(13482..13694)
/rpt_family="AluDb"
complement(13709..14037)
/rpt_family="L1"
complement(14009..14105)
/rpt_family="L1MC/D"
complement(14115..14415)
/rpt_family="AluY"
complement(14417..14849)
/rpt_family="L1MC/D"
complement(14859..15153)
/rpt_family="AluSx"
15226..15492
/rpt_family="AluJo"
complement(16257..16340)
/Note="Predicted exon, program: g1all2exons_human_1.3,

```

```

repeat_region /rpt_family="AluSg"
16362..16666
complement(16691..17060)
/rpt_family="THEIC"
repeat_region /rpt_family="AluJo"
17482..17781
/rpt_family="L1"
complement(17782..17870)
/rpt_family="(GAA)n"
17968..18273
/rpt_family="AluSg"
complement(18401..18817)
/rpt_family="L1"
complement(18818..19193)
/rpt_family="HERVH"
complement(19194..19412)
/rpt_family="HERVH"
complement(19414..20028)
/rpt_family="HERVH"
complement(20031..23647)
/rpt_family="HERVH"
complement(23651..24066)
/rpt_family="L1"
complement(24403..24714)
/rpt_family="HERV4"
24789..25080
/rpt_family="AluSx"
25161..25207
/rpt_family="L1MD3"
25224..25509
/rpt_family="AluDb"
25605..25901
/rpt_family="AluSg"
25902..26069
/rpt_family="AluSg/x"
complement(26224..26256)
/rpt_family="(TAA)n"
complement(26257..26545)
/rpt_family="AluJo"
26573..26876
/rpt_family="AluJo"
complement(27125..27422)
/rpt_family="AluSx"
27866..28173
/rpt_family="AluSx"
28181..28379
/rpt_family="AluSg"
complement(28359..28522)
/rpt_family="AluSg"
complement(28526..28705)
/rpt_family="AluSg"
complement(28733..28872)
/rpt_family="AluDb"
28874..29081
/rpt_family="AluJo"
29463..29583
/Note="Predicted exon, program: g1all2exons_human_1.3,
frame: 2, quality: good, score: 71.000"
repeat_region /rpt_family="AluSg"
30336..30634
/rpt_family="AluSg"
30646..30779
/rpt_family="AluDb"
complement(30830..31040)
/rpt_family="L1"

```

```

Alignment Scores:
Pred. No.: 7.62e-19
Score: 289.00
Percent Similarity: 48.01%
Best Local Similarity: 45.51%
Query Match: 43.26%
Length: 146180
Matches: 71
Conservative: 4
Mismatch: 30
Indels: 53

```

Dd	23	AGCATGAGGCCACGAATGGTCGACTTGTGGCTCTTGCTGCTGGATAACAGTTCTCCCTC	82
Oy	27	LeucysSerHISGLYALAProValAlAPromethrProTYrLEUmetLeuCysInPro	46
Dd	83	CAGTGTCAAAAGAAGAACACAGACCGCTCTGTGGCTCAGACACTGTGGCTGTGCCACGC	142
Oy	47	HISLysArgCysgSLyAsPlySPheTYrAsPrProleuGlnHISCysCYSTyrASpAsPaLA	66
Dd	143	ACACCACAGGTGTGGGAACAAGATCTACAACCCCTTCAGAGCAGTGTGTATTGATGATGCC	202
Oy	67	valValProlEuaAlAargThGlnThCysgSLyasn--CystHrPheArVAlCysPhe	85
Dd	203	ATCTTAATCTTAAAGAGAACCCGCCGCTGTGGTCCACCTCACCTTCTGGCCCTGCTT	262
Oy	86	GIUGlIncCysCySProtYPthr-----PhemETAllysLeu-----	97
Dd	263	GAGCTCTGCTGCCGAGCTTTTGGCCCCCAGCAGAAAGTTCTTGTGAAGTTAGGGTT	322
Oy	98	-----IleaSglnAsnCysAspSeraLaArg	106
Dd	323	CTGGTATGAGAGTCACTGTCACCTTATCTCCCACTCTCCGAGACTGTACAGGACAGG	382
RESULT 6			
LOCUS	AX055696	662 bp	DNA linear PAR 13-JAN-2001
DEFINITION	Sequence 11 from Patent WO0073348.		
ACCESSION	AX055696		
VERSION	AX055696.1 GI:12228831		
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homniidae; Homo. Baker,K.P., Goddard,A., Gurney,A.L., Hebert,C., Henzel,W., Kabatnik,R.C., Shelton,D.L., Smith,V., Watanabe,C.K. and Wood,W.I. Methods and compositions for inhibiting neoplastic cell growth Patent: WO 0073348-A 11 07-DEC-2000.		
TITLE			
JOURNAL	Genentech, Inc. (US)		
FEATURES			
source	Location/Qualifiers		
	1..662		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	152 a 187 c 167 g 156 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	6,74e-16	Length:	662
Score:	237.00	Matches:	49
Percent Similarity:	55.00%	Conservative:	17
Best Local Similarity:	40.83%	Mismatches:	34
Query Match:	35.48%	Indels:	20
GB:	6	Gaps:	3
US-09-092-297-17 (1-117) x AX055696 (1-662)			
Oy	7	AlaMetAlAProAargGLyCysILeValAlAVALPheAlAliePheCysIIeserArLeu	26
Dd	23	AGCATGAGGCCACGATGTCATCTTGGCTTCTGCTGGATACAGTTCCTCCTC	82
Oy	27	LeucysSerHISGLYALAProValAlAPromethrProTYrLEUmetLeuCysInPro	46
Dd	83	CAGTGTCAAAAGAAGAACACAGACCGCTCTGTGGCTCAGACACTGTGGCTGTGCCACGC	142
Oy	47	HISLysArgCysgSLyAsPlySPheTYrAsPrProleuGlnHISCysCYSTyrASpAsPaLA	66
Dd	143	ACACCACAGGTGTGGGAACAAGATCTACAACCCCTTCAGAGCAGTGTGTATTGATGATGCC	202
Oy	67	valValProlEuaAlAargThGlnThCysgSLyasn--CystHrPheArVAlCysPhe	85
Dd	203	ATCTTAATCTTAAAGAGAACCCGCCGCTGTGGTCCACCTCACCTTCTGGCCCTGCTT	262

Qy	86	GIU6GACGCGCyscysprotrprrphr-----	Phemervallysleu-----	97
Db	263	GAAGTCTGCTGTCGCGCAGACCTTTTGGCCCCCAGACAGAAATCTTGTGTGAAGTTGAGGCTT	:	322
Qy	98	-----	:	106
Db	323	CTGGGTATCAAGTCTCAGCTCACTTATCTCCCACTCTCCGAGAGCTGTACACAGAAACAGG	:	382
RESULT 7				
LOCUS	AX118905			
DEFINITION	Sequence 69 from Patent WO0129221.			
ACCESSION	AX118905			
VERSION	AX118905.1			
KEYWORDS	GI:14035859			
SOURCE				
ORGANISM	human.			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
JOURNAL	1 (bases 1 to 378)			
FEATURES	Conklin,D.C. and Yee,D.P.			
source	Proteins and polynucleotides encoding them			
	Patent: WO 0129221-A 69-26-APR-2001;			
	ZymoGenetics, Inc. (US)			
	Location/Qualifiers			
	1..378			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	1..378			
	/note="unamed protein product"			
	/codon_start=1			
	/protein_id="CA038526.1"			
	/db_xref="GI:14035860"			
	/translation="MPRPCCIALVCMITVFLLOCSKGTDAVPAGSLIMCOPTPRCG			
	NKIYNPEOCYDAILSLKETRRCGSGTCFWMPCFELCDESFSPQOKFLVKLVLMG			
	KSQCHLSPISCSCTRRNRHLYP"			
	BASE COUNT 76 a 108 c 95 g 99 t			
ORIGIN				
Alignment Scores:				
Pred. No.:	4.73e-16	Length:	378	
Score:	236.00	Matches:	49	
Percent Similarity:	54.62%	Conservative:	16	
Best local Similarity:	41.18%	Mismatches:	34	
Query Match:	35.33%	Indels:	20	
Db:	6	Gaps:	3	
US-09-092-297-17 (1-117) x AX118905 (1-378)				
Qy	8	MetaLAProAgGlyCysIleValAlaValAlaIlePheCysIleSerArgLeu	27	
Db	1	ATGAGGCCACATGCTGATCTTGGCTTTCTGCTGGATTAACAGTCTTCTCCACAG	60	
Yy	28	CysSerHisGlyAlaProValAlaProMetThrProTyrIleMetLeuCysGlnProHis	47	
Db	61	TGTTCAAAAGGAATACAGACGCTCTGTTGGTGCACAGAGCTGGCTGCACACCCACA	120	
Qy	48	LysArgCysGlyAspLysPheTyrAspProLeuGlnHisCysCysTyrAspAspAlaVal	67	
Db	121	CCCAAGTGTGGAAACAAGATCTACACCCCTTCACAGACAGTCTCTTATGATGATGACCAT	180	
Qy	68	ValProLeuAlaAgtInGlnInThCysGlyAsn---CysThrPheArgValCysPheGlu	86	
Db	181	TTATCCCTTAAGAGAGACCCGCCGCTGTGGCTCCACCTTGACCTTCTGGCCCTGCTTGG	240	
Qy	87	GlnCysCysProTrrprrh-----PheMetValLysLeu-----	97	
Db	241	CTCTGCTGTCCCGAGATCTTTTGGCCCCCAGACAGAAATCTTGTGAAGTTGAGGCTTCTG	300	
Qy	98	-----IleAsnIleAsnCysAspSerAlaArg	106	
Db	301	GGTATGAAGTCTCAGTGTCACTTATCTCCCACTCTCCCGGAGCTGTACACAGAAACAGG	357	

RESULT 8
AC007785/C
DEFINITION AC007785 166500 bp DNA linear PRI 11-JUN-1999
AC007785 Homo sapiens chromosome 19, BAC 282485 (CIT-B-344H19), complete
ACCESSION AC007785
VERSION AC007785.1 GI:5042403
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Lamerdin,J.E., McCreedy,P.M., Richardson,P., Sakladasis,G.,
1 (bases 1 to 166500)
Lamerdin,J.E., Gordon,L., Scott,D., Johnson,G., Stillwagen,S.,
Burkhardt-Schultz,K., Gordon,L., Regala,W., Terry,A., Dangnan,L.,
Phan,H., Velasco,N., Do,L., Regala,W., Terry,A., Dangnan,L.,
Erler,A., Christensen,M., Georgescu,A., Avila,J., Atlix,C.,
Andruse,T., Amico-Keller,G., Coefield,J., Duarte,S., Lucas,S.,
Baird,R., Thomas,P., Quan,G., Kronmiller,B., Arellano,A.,
Sanders,C., Ow,D., Nolan,M., Trong,S., Kobayashi,A., Olsen,A.S. and
Carrano,A.V.
TITLE Sequence analysis of a 2.5 Mb. region in 19q13.2-13.3 between APOE
and C5R1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 166500)
AUTHORS Lamerdin,J.E.
TITLE Direct Submission
JOURNAL Submitted (11-JUN-1999) Joint Genome Institute, Lawrence Livermore
National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
Map and sequence oriented from 9 centimorgans to telomere, BAC 282485
overlaps cosmid R30477 to the left from chromosome 19, 4,419 of this
accession, and overlaps BAC 264576 (CIT-B-297N14, AC006262) from
bases 146,203 to 166,500. Additional chromosome 19 map and sequence
information may be obtained at:
http://www.bio.liln.gov/bdpr/genome/genome.html.
FEATURES
Source Location/Qualifiers
1. 166500
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.3 between DM and C5R1"
/clone="CIT-B-344H19"
/clone_1lb="Cal Tech CIT-B BAC library"
/note="LILN clone name: BC282485"
70..145
/rpt_family="(TGAA)n"
complement(177..435)
/rpt_family="Aluub"
complement(472..712)
/rpt_family="Mir"
746..904
/note="DPS similarity to overlapping ESTs:
N37722 yx92e04.r1 Homo sapiens cDNA clone 269214 5';
(12..1170); 99% identity..AA283147 z14d07.r1 NCI-CCAP_GCB3
Homo sapiens cDNA clone IMAGE:713101 5'; (6..163); 99%
identity..AA143160 z149d07.r1 Soares pregnant uterus NDBPU
Homo sapiens cDNA clone 505261 5'; (1..43); 100%
identity..AA142861 z149d07.s1 Soares pregnant uterus NDBPU
Homo sapiens cDNA clone 505261 3'; (570..527); 64%
identity."
join(<746..904,1157..1239,4294..4451,5875..6036,
12784..12994,13080..13162,13903..13958,14053..14128,
14429..14568,14726..14804,14982..15039,15813..15885,
16025..16122)
/note="Hypothetical arginine-rich gene product;
Hypothetical CDS constructed from overlapping ESTs and
Xiral predictions. EST coverage is lacking for some
portions of the CDS. Gene may be alternatively spliced, as
multiple transcripts map into this region; alternatively,
these individual clones may represent artifactual events
during RT and cDNA creation"
/codon_start=3
/evidence=not_experimental

```

/product="BC282485.1"
/protein_id="AAD38244.1"
/db_xref="GI:5042405"
/translat="TAAADGAGAGQVDFYGVGEHAAVAVASGVLELEVEDRMTADQ
WGEEDAFIEDLTKTGNFKQFNIPTCHMIESALTOSESVTDLLTYDLESLARK
MGGRGSLGSPNSAONSKRYLILTYSEPRINRPLRPOGKPDVVLGGITRSKE
ELGRGLGADSCONTDRTRENEHMLREOVSVLSASKRELEAOJRSREDEALAGRAFOE
AVALAGLVRCGLLELDRBRGLGHVAVARGQDCCRKLAKLEEKASERISLAKRLATL
LSLALYKGGKSGRGRRPAPSPPTGGALRPDPAFKAERQREISLAKRLATL
RFGSGSDSDGSVSGRQTPALPTGRGAPAPNSNRKSVKSPFSSRCSASCGL
EDFSLRSRGHGRRGKRPSPPTPGSGNSKSPVERSHHQKSLANGSGWVPIKESSE
HQAAMARLIDRLAKALAEQYMRDLMS"
misc_feature
1157..1239
/note="Predicted exon, program: grail2exons_human_1.3,
frame: 0, quality: good, score: 62.000-DDS similarity to
overlapping ESTs:
N39722.yx92e04.r1 Homo sapiens cDNA clone 269214 5';
(171..255): 95% identity.-AA283147.zt14d07.r1
NCI-CGAP-CGB1 Homo sapiens cDNA
clone IMAGE:713101 5'; (164..247); 92% identity.-AA143160
z149d07.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA
clone 505261 5'; (44..126): 100% identity.-N28707
yx66d11.r1 Homo sapiens cDNA clone 266709 5'; (1..66): 98%
identity.-AA142881.z149d07.s1 Soares pregnant uterus NbHPU
Homo sapiens cDNA clone 505261 3'; (526..438): 82%
identity.
1536..1634
/rpt_family="CA.n"
/complement(1673..1973)
/rpt_family="Alusx"
/complement(2040..2185)
/rpt_family="LINE2"
/complement(2200..2498)
/rpt_family="Alusx"
/complement(2576..2746)
/rpt_family="Alusg/x"
/complement(2747..3044)
/rpt_family="Alusx"
3048..3750
/rpt_family="MER64"
3841..4060
/rpt_family="MER64"
/rpt_family="MER64"
4294..4451
/note="DDS similarity to overlapping ESTs:
N39722.yx92e04.r1 Homo sapiens cDNA clone 269214 5';
(256..438): 81% identity.-AA283147.zt14d07.r1
NCI-CGAP-CGB1 Homo sapiens cDNA
clone IMAGE:713101 5'; (248..400); 99% identity.-AA143160
z149d07.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA
clone 505261 5'; (127..281): 97% identity.-N28707
yx66d11.r1 Homo sapiens cDNA clone 266709 5'; (67..222);
98% identity.-AA142881.z149d07.s1 Soares pregnant uterus
NbHPU Homo sapiens cDNA clone 505261 3'; (437..275); 97%
identity.
complement(4789..5083)
/rpt_family="Alusx"
5875..5898
/note="DDS similarity to N28707.yx66d11.r1 Homo sapiens
cDNA clone 266709 5'; (223..246); 100% identity."
5875..5972
/note="DDS similarity to overlapping ESTs:
N39722.yx92e04.r1 Homo sapiens cDNA clone 269214 5';
(433..506): 71% identity.-AA143160.z149d07.r1 Soares
pregnant uterus NbHPU Homo sapiens cDNA clone 505261 5';
(282..379); 100% identity.-AA142881.z149d07.s1 Soares
pregnant uterus NbHPU Homo sapiens cDNA clone 505261 3';
(274..175): 95% identity."
5906..6036
/note="Predicted exon, program: grail2exons_human_1.3,
frame: 1, quality: excellent, score: 100.000"
complement(6148..6443)
/rpt_family="Alusx"
complement(6608..6648)

```

```

repeat_region      /rpt_family="LINE2"
                   complement(7118..7340)
repeat_region      /rpt_family="MIR"
                   complement(7538..7639)
repeat_region      /rpt_family="(CA)n"
                   complement(7792..8054)
repeat_region      /rpt_family="MIR"
                   complement(8789..9083)
repeat_region      /rpt_family="Aluuc"
                   9456..9520
repeat_region      /rpt_family="MER5B"
                   complement(9539..9841)
repeat_region      /rpt_family="Alusx"
                   9909..9944
repeat_region      /rpt_family="MER5B"
                   10099..10401
repeat_region      /rpt_family="Alusx"
                   complement(10587..10752)
repeat_region      /rpt_family="Aluv"
                   complement(10753..11057)
repeat_region      /rpt_family="Alusx"
                   complement(11065..11191)
repeat_region      /rpt_family="Aluv"
                   11322..11619
repeat_region      /rpt_family="Alusg"
                   11775..11814
repeat_region      /rpt_family="AT_rich"
                   complement(11864..11885)
repeat_region      /rpt_family="AT_rich"
                   complement(11914..11976)
repeat_region      /rpt_family="Alus"
                   complement(12055..12088)
repeat_region      /rpt_family="AT_rich"
                   complement(12186..12327)
repeat_region      /rpt_family="L1M1"
                   complement(12334..12384)
repeat_region      /rpt_family="LINE2"
                   12784..12994
misc_feature        /note="predicted exon, program: gail2exons human_1.3,
                   frame: 1, quality: excellent, score: 92.000"
repeat_region      12847..12808
/rpt_family="GC_rich"
misc_feature        12895..12994
/note="SADS similarity to AA765166 n279c05.s1 NCI CGAP GCBI
                   Home sapiens cDNA clone IMAGE:1301672; (222..123); 97%
                   identity."
repeat_region      13002..13032
/rpt_family="GC_rich"
misc_feature        13080..13162
/note="Predicted exon, program: gail2exons human_1.3,
                   frame: 2, quality: excellent, score: 100.000-DPS
                   similarity to AA765166 n279c05.s1 NCI CGAP GCBI Homo
                   sapiens cDNA clone IMAGE:1301672; (122..48); 100%
                   identity."
repeat_region      complement(13799..13954)
/rpt_family="GC_rich"

Alignment Scores:
Pred. NO.:         3.3e-09          length:      166500
Score:             197.50           Matches:      37
Percent Similarity: 65.28%          Conservative: 10
Best Local Similarity: 51.3%        Mismatches:   18
Query Match:       29.57%          Indels:       7
DB:                9                Gaps:         2

US-092-297-17 (1-117) x AC007785 (1-166500)

OY      33 ProValAlaProMetIhPrOTytleuMetleuCysGlnProHnSLySAtgCysGlyAsp 52
      ||| |||||::: ||| |||||::: ||| |||||::: ||| |||||:::
Db 121894 CCAAGACGCTCTGTGGCTCAGAGACTGTGGCTGCCAGCACCCAGCCAGGTGTGGGAAAC 121835

53 LysPheTyTAspRProLeuGlnHnIScysGtTyTAspRAspAlaValAlaProLeuAlaArg 72
||| |||||::: ||| |||||::: ||| |||||::: ||| |||||:::

```

Db	121834	AAGATGTACAAACCCTTCANAGAGAGTCCGCTTTATGATNGATGCATCTTATTCCTTAAGAAGC	121775
OY	73	ThrlnJnthrCyseGlyAsn---CystThrpheargValCysPheglucInlcYscypProtip	91
Db	121774	ACCCGCCCGCTGGCTGCACCTGCACCTTGCCGCTGGCTTTGAGGTCTGCGTGTCCGAG	121715
OY	92	Thr-----PheMetValylsLeu	97
Db	121714	TCTTTTGGCCCCAGACAGATTCTCTGTGAAGTTG	121679
RESULT 9			
LOCUS	AC007785	Homo sapiens chromosome 19, BAC 282485 (CIT-B-344H19), complete	
DEFINITION	Sequence.		
ACCESSION	AC007785		
VERSION	AC007785.1	GI:5042403	
KEYWORDS	HTG.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 166500) lamerdin,J.E., McCreedy,P.M., Richardson,P., Sakadatsis,G., Burkhardt-Schultz,K., Gordon,L., Scott,D., Johnson,G., Seilwagen,S., Phun,H., Velasco,N., Do,L., Regala,W., Terry,A., Dangnan,L., Etler,A., Christensen,M., Georgescu,A., Avila,J., Attix,C., Andreise,T., Amico-Keller,G., Coefield,J., Duarte,S., Lucas,S., Bruce,R., Thomas,P., Quan,G., Kronmiller,B., Arellano,A., Sanders,C., Ow,D., Nolan,M., Trong,S., Kobayashi,A., Olsen,A.S. and Carreno,A.V.		
TITLE	Sequence analysis of a 2.5 kb region in 19q13.2-13.3 between APOE and CSRI		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 166500)		
AUTHORS	Lamerdin,J.E.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-JUN-1999) Joint Genome Institute, Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94551, USA Map and sequence oriented from q centromere to telomere. BAC 282485 overlaps cosmid R30477 to the left from bases 1 to 4,419 of this accession, and overlaps BAC 264576 (CIT-B-297N14, AC006262) from bases 146,203 to 166,500. Additional chromosome 19 map and sequence information may be obtained at: http://www.bio.1nl.gov/bdrr/genome/genome.html .		
COMMENT	Location/Qualifiers		
FEATURES	1. 166500		
SOURCE	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="19"		
	/map="19q13.3 between DM and CSRI"		
	/clone="CIT-B-344H19"		
	/clone_lib="Cal Tech CIT-B BAC library"		
	/note="L1NL clone name: BC282485"		
	70..145		
	/rpt_family="(TGA)n"		
	complement(17)..(453)		
	/rpt_family="Alub"		
	complement(472..712)		
	/rpt_family="MR"		
	746..904		
	/note="DSS similarity to overlapping ESTs: N9722 yx92e04.r1 Homo sapiens cDNA clone 269214 5'; (12..170); 99% identity.-AA283147 zti4d07.r1 NCI CGRP GCB Homo sapiens cDNA clone IMAGE:713101 5'; (6..163); 99% identity.-AA143160 z149d07.r1 Soares pregnant uterus Nbhpv Homo sapiens cDNA clone S05261 5'; (1..43); 100% identity.-AA142881 z149d07.s1 Soares pregnant uterus Nbhpv Homo sapiens cDNA clone S05261 3'; (570..527); 64% identity."		
	j01n<746..904,1157..1239,4294..4451,5875..6036, 12784..12994,13080..13162,13903..13958,14053..14128, 14429..14568,14726..14804,14982..15039,15813..15885,		

```

16025..16122)
/hypothesis="Hypothetical arginine-rich gene product;
Hypoetical CDS constructed from overlapping ESTs and
Xgail predictions. EST coverage is lacking for some
portions of the CDS. Gene may be alternatively spliced,
multiple transcripts map into this region; alternatively,
these individual clones may represent artifactual events
during RT and cDNA creation"
/codon_start=3
/evidence=not_experimental
/product="BC82485.1"
/protein_id="AA03824.1"
/db_xref="GI:5042405"
/translacion="TLMAGPGLOVDYVFRGVHNAVMSGVLELEDYEDRWTAQO
WRESEFAGIEDLTHRTGNFQKQFNIFCHMLSAALQSSSEVTLDTLYTDESIRNKK
MGRPSPLSRASQNLNRSKRYLILITYSVEDRHWPLPLPYGGRDPVIOGIIIRAKE
ELRGRLDGLGMDRDTRENEIHWLREOVSLRSEKRELELOGSRSEALAGRIARKE
AALRLGVGLLELELEROEGSLRHVARGGDCDCLAKELDEARASERSLRARITVT
SELSGALRGGRSRRGRGPRAPSPSPGGRALRDPFAFYVAKRRKRELOMQOONRN
RUSGSGGSGGPGSVWSRPTOPPALTLTGDDAPNRSRNSSVDSEFRSCSSASCSPL
EDPSELSGGRHRRKRGKPPSPYPMGSMKSPVSRSHHOKSLANSGWAPIKEYSEE
HQADEAEIDEDRLARKLQEFWMRLDQRS"
1157..1239
/notes="predicted exon, program: gtail2exons_human.1.3,
frame: 0, quality: good, score: 62.000-DDS similarity to
overlapping ESTs:
N39722.yx92e04.r1 Homo sapiens cDNA clone 269214 5';
171..255); 95% identity.-AA283147 zt14d07.r1
NCI CGAP GCBI Homo sapiens cDNA
clone IMAGE:713101 5'; (164..247); 92% identity.-AA143160
z149d07.r1 Soares pregnant uterus NbHPV Homo sapiens cDNA
clone 505261 5'; (44..126); 100% identity.-N28707
yx66d11.r1 Homo sapiens cDNA clone 266709 5'; (1..66); 98%
identity.-AA142881 z149d07.s1 Soares pregnant uterus NbHPV
Homo sapiens cDNA clone 505261 3'; (526..438); 82%
identity."
1536..1634
/rpt_family="CA)n"
complement(1673..1973)
/rpt_family="Alusx"
complement(2040..2155)
/rpt_family="LINE2"
complement(2200..2498)
/rpt_family="Alusx"
complement(2516..2746)
/rpt_family="Alusg/x"
complement(2747..3044)
/rpt_family="Alusx"
3048..3750
/rpt_family="MER64"
3841..4060
/rpt_family="MER64"
4294..4451
/notes="BDS similarity to overlapping ESTs:
N39722.yx92e04.r1 Homo sapiens cDNA clone 269214 5';
(256..438); 81% identity.-AA283147 zt14d07.r1
NCI CGAP GCBI Homo sapiens cDNA
clone IMAGE:713101 5'; (248..400); 99% identity.-AA143160
z149d07.r1 Soares pregnant uterus NbHPV Homo sapiens cDNA
clone 505261 5'; (127..281); 97% identity.-N28707
yx66d11.r1 Homo sapiens cDNA clone 266709 5'; (67..222);
98% identity.-AA142881 z149d07.s1 Soares pregnant uterus
NbHPV Homo sapiens cDNA clone 505261 3'; (437..275); 97%
identity."
complement(4789..5083)
/rpt_family="Alusx"
5875..5898
/notes="BDS similarity to N28707.yx66d11.r1 Homo sapiens
cDNA clone 266709 5'; (223..246); 100% identity."
5875..5972
/notes="BDS similarity to overlapping ESTs:
N39722.yx92e04.r1 Homo sapiens cDNA clone 269214 5';
(433..506); 71% identity.-AA143160 z149d07.r1 Soares

```

DB: 9 Gaps: 5

US-09-092-297-17 (1-117) x AC007785 (1-166500)

OY	22	CysIISeSerArgLeuLecCysSserHisGlyAlaProValAlaIAPrometThrProTyrIleu	41
Db	158329	TGTCTGTC-----ATCTGTCCA-----	
OY	42	MettUeCysGlnPronHISLySarGCySLypSyPheryrAspProLeuGlnHisCys	61
Db	158368	TGGCTGTGCCAGCCCGCCAGCCAGGTGTGGAGAACAAGACTCAACACCCTTGAGACAGTGC	158427
OY	62	CysTrjAspARPAIvalAvalProleuHlaArghrInThrysgly---Asmcsythr	80
Db	158428	TGTTACATGTAGCGCATCTGTCTCCCTGTAGCACAGACCAGCAATTGTGTCCCCCTGTGCAC	158487
OY	81	PheArGVAlCySPhegluglIncycCs-----ProtRphrPhemet	94
Db	158488	TTCTGGCCCTCTGTTAGAGCTCTGGCTGTCTATTCTTGCCCTCCACAAAGATTTTCTT	158547
OY	95	VallyslEu-----IlEasnglInsncysApsserAlarTyhSer	108
Db	158548	GTTGAAGCTGAAGGTTCAAGGTTGATTAATTCACAGTCACATCATCTCCATCTCCC	158601

RESULT 10	AC094931	47407 bp	DNA	linear	HTG 20-DEC-2001	
LOCUS	AC094931/c					
DEFINITION	Rattus norvegicus clone CH230-6D18,	*** SEQUENCING IN PROGRESS.	***,			
ACCESSION	AC094931					
VERSION	AC094931.2					
KEYWORDS	GI:17941731					
SOURCE	HTG; HTGS PHASE I.					
ORGANISM	Norway rat. Rattus norvegicus Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
REFERENCE	1 (bases 1 to 47407) Mizuy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Aishbrook,S.L., Amaralunga,H.C., Aye,U.J.R., Baks,T.T., Barbaila,J., Benton,J., Blincke,K., Blankenburg,K., Bonnin,D., Bouck,J., Beaton,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhaý,C., Burich,P., Burkelt,C., Burrell,K.L., Byrd,N.C., Carton,T.F., Carier,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Demn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,D., Foster,P., Frantz,P., Gabisi,A., Gao,Y., Garcia,A., Garner,T., Garza,N., Gall,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homsl,F., Howard,S., Huber,J., Halys,S., Hume,J., Jacksom,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolyvet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Kovvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichteage,O., Lieu,C., Liu,Q., Liu,W., Loulseng,H., Lozaro,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,U., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Meil,G., Melker,M., Miner,G., Miner,Z., Mitchell,T., Mohabat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newsham,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkenwo,S., Oguh,M., Okunnu,G., Orangunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Plums,E., Pu,L.L., Qualls,M., Ren,Y., Rivers,M., Rojas,A., Rojubokhan,I., Rolfe,M., Ruiz,S., Savory,G., Scherer,S., Scott,G., Shen,H., Shooshkari,N., Slison,I., Sodergren,E., Sonalik,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tanerisa,A., Tanerisa,K., Tang,H., Taney,J., Taylor,C., Taylor,C., Taylor,T., Teiford,B., Thomas,N., Thomas,S., Usmani,K., Vasconcelos,V., Villalba,B., Williams,N.,					

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
Wallington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S.,
Moriarty, R., Wu, C., Wu, Y., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.
Unpublished
2 (bases 1 to 47407)

COMMENT

Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:15624768.

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GRH
Center clone name: CH230-6D18
Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findphraplist

Consensus quality: 22954 bases at least Q40
Consensus quality: 29327 bases at least Q30
Consensus quality: 35153 bases at least Q20
Estimated insert size: 22926; sum-of-coverage estimation
Quality coverage: 0x in Q20 bases; average-CP estimation
Quality coverage: 0.2x in Q20 bases; sum-of-coverage estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1945: contig of 1945 bp in length
* 1946 2045: gap of unknown length
* 2045 4298: contig of 2253 bp in length
* 4299 4398: gap of unknown length
* 4399 6397: contig of 1999 bp in length
* 6398 6498: gap of unknown length
* 6498 8809: contig of 2312 bp in length
* 8810 8909: gap of unknown length
* 8910 11463: contig of 2554 bp in length
* 11464 11563: gap of unknown length
* 11564 13601: contig of 2038 bp in length
* 13602 14916: contig of 1215 bp in length
* 14917 15016: gap of unknown length
* 15017 17039: contig of 2023 bp in length
* 17040 17139: gap of unknown length
* 17140 18161: contig of 1477 bp in length
* 18161 18716: gap of unknown length
* 18717 20718: contig of 2001 bp in length
* 20718 20818: gap of unknown length
* 20818 22641: contig of 1824 bp in length
* 22642 22741: gap of unknown length
* 22742 24505: contig of 1764 bp in length
* 24506 26758: gap of unknown length
* 26759 26858: gap of unknown length
* 26859 27884: contig of 1025 bp in length
* 27884 27984: gap of unknown length
* 27984 29980: contig of 1997 bp in length
* 29981 31208: gap of unknown length
* 31209 31308: gap of unknown length

FEATURES
source
1. 47407
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-6D18"

BASE COUNT 12893 a 9169 c 10001 g 12470 t 2874 others

ORIGIN

Alignment Scores:
Pred. No.: 52.1
Score: 94.50
Percent Similarity: 32.418
Best Local Similarity: 24.078
Query Match: 14.154
DB: 2
Gaps: 3

US-09-092-297-17 (1-117) x AC094931 (1-47407)

QY 1 ProLeuGlnProFroArgAlaMetaLapProArgGlyCysIleValAlaValPheAlaIle 20
DB 35741 CCCCTTACTGCCCCCTCCCTGTTGCCCCCTTACTGCCCCCTTGTGTGACGCCCT 35682
QY 21 PheCysIleSer-----ArgLeuLeuCysSerHisGlyAlaPro 33
DB 35681 TTCTGCACCTCGTGTGTGCACACACCTTACCTGCTGTTGCACACACCTTACCT 35622
QY 34 ValAlaProMetThrProTyrIleuMetLeuGlyGlnProHisLysArgGlySplys 53
DB 35621 GCCCCTCCGTTGTC-----ACACACCTTACCTGACCTGCTCT 35583
QY 54 PheTyrAspProLeuGlnHisCysCysTyrAspAspAlaVal---ValProLeuAlaArg 72
DB 35582 TGTACACACCTTACTGCTGCTGTTGTGCACACACCTTACCTGACACCTTACCTGAC 35523
QY 73 ThrGlnThrCysGlyAspGlySerThrPheArgValCysPheGlnGlnCysCysProTyrPhe 92
DB 35522 ACACCTTATCTATCTCCATGATATCTGTGACAGCTGTTTAAAGAGCTGCGCCCTGCTGT 35463
QY 93 PheMetValIleLeuIleAsnGln 100
DB 35462 GGTCTGCACAGTCTGCTGTCACAGAA 35439

RESULT 11
AC094152/c 115627 bp DNA linear HTG 20-DEC-2001
LOCUS AC094152 Rattus norvegicus clone CH230-2122, *** SEQUENCING IN PROGRESS ***
DEFINITION 61 unordered pieces.
ACCESSION AC094152.2 GI:17940864
VERSION AC094152.2
KEYWORDS HTG; HTGS_PHASE1.

SOURCE
ORGANISM
NORWAY rat.
Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
AUTHORS
1 (bases 1 to 115627)
Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C.,
Alstbrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbarella,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonin,D., Boucek,J.,
Bowles,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Bulay,C.,
Burgh,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyte,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Denny-Carrill,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Dem,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Gatta,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Joliver,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,O., King,L., Koryah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewised,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mathney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,R.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenko,S.,
Oguchi,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojudoxan,T., Rolfe,M.,
Rutz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shiohatai,N.,
Slison,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Washington,S., Williams,G., Williamson,A., Wleczky,R., Woodan,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gbbs,R.

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
COMMENT
Unpublished
2 (bases 1 to 115627)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:115636085.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GACA
Center clone name: CH230-2L22
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhraplist
Consensus quality: 93124 bases at least Q40
Consensus quality: 101400 bases at least Q30
Consensus quality: 107250 bases at least Q20
Estimated insert size: 81290; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 1.1x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 61 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1863	contig of 1863 bp in length
1864	1963	gap of unknown length
1964	4931	contig of 2968 bp in length
4932	5031	gap of unknown length
5032	7168	contig of 2137 bp in length
7169	7268	gap of unknown length
7269	10284	contig of 3015 bp in length
10284	10383	gap of unknown length
10384	12877	contig of 2494 bp in length
12878	12977	gap of unknown length
12978	16257	contig of 3280 bp in length
16258	16357	gap of unknown length
16358	20013	contig of 3656 bp in length
20014	20113	gap of unknown length
20114	23261	contig of 3148 bp in length
23262	23361	gap of unknown length
23362	25699	contig of 2338 bp in length
25700	25799	gap of unknown length
25800	27755	contig of 1956 bp in length
27756	27855	gap of unknown length
27856	29751	contig of 1896 bp in length
29752	29851	gap of unknown length
29852	32021	contig of 2170 bp in length
32022	32121	gap of unknown length
32122	34707	contig of 2386 bp in length
34708	34807	gap of unknown length
34808	37259	contig of 2452 bp in length
37260	37359	gap of unknown length
37360	39491	contig of 2132 bp in length
39492	39591	gap of unknown length
39592	41358	contig of 1767 bp in length
41359	41588	gap of unknown length
41590	43614	contig of 2156 bp in length
43615	43714	gap of unknown length
43715	46666	contig of 2952 bp in length
46667	46766	gap of unknown length
46767	48026	contig of 1260 bp in length
48027	48126	gap of unknown length
48127	50299	contig of 2173 bp in length
50300	50399	gap of unknown length
50400	52442	contig of 2043 bp in length
52443	52542	gap of unknown length
52543	54526	contig of 1884 bp in length
54527	54626	gap of unknown length
54627	56327	contig of 1601 bp in length
56328	56327	gap of unknown length
56328	58509	contig of 2182 bp in length
58510	58609	gap of unknown length
58610	60960	contig of 2251 bp in length
60961	62366	gap of unknown length
62367	62466	gap of unknown length
62467	63856	contig of 1390 bp in length
63857	63956	gap of unknown length
63957	66010	contig of 2054 bp in length
66011	66110	gap of unknown length
66111	68622	contig of 2512 bp in length
68623	68722	gap of unknown length
68723	69821	contig of 1099 bp in length
69822	69921	gap of unknown length
69922	71830	contig of 1909 bp in length
71831	71930	gap of unknown length
71931	73224	contig of 1294 bp in length
73225	73324	gap of unknown length

BASE COUNT 44657 a 34364 c 33282 g 45690 t 332 others
ORIGIN

Alignment Scores:

Pred. No.: 437
Score: 91.00 Length: 158325
Percent Similarity: 40.97% Matches: 40
Best Local Similarity: 27.78% Mismatches: 19
Query Match: 13.62% Indels: 47
DB: 2 Gaps: 38

US-09-092-297-17 (1-117) x AC087564 (1-158325)

```

Oy 2 LeuGlnProPro-----5
Db 36796 CTGCACACCTCCACCTCGGGTCAAGCTATTCTCCCTCAGCCTCTGAGTACGCTGG 36737
Oy 6 -----ATGAlaMetAlaProArgGlyCysIleValAlaAlaPheAlaIlePheCysIle 23
Db 36736 GACTACAGTGCATACACACAGCCAGTATATATTATT-----ATTTTGTGATTT 36683
Oy 24 ---SerArgLeuLeuCysSerHisGlyAlaProValAlaProMetThrProTyrLeuMet 42
Db 36682 TTTAGTACAGACGTGGTTTCCACCATGTTGGCCAGGATGTCATCTCTGACCTCGCTG 36623
Oy 43 LeuCysGlnPro-----HisLysArgCysGlyAspLysPheTyrAspProLeuGln-----59
Db 36622 ATCGCGCCGCTTAGCCGCCCAAGTCCGGGATTTATAGCGCTGAGCCCTCAGCCTGAGCC 36563
Oy 60 HisCysCysTyr---AspAlaValAlaProLeuAlaArgThrGlnThrCysGlyAsn 78
Db 36562 TATGCTGCTATTTCACACATGATATTAACCTTCGATCTT-----AGT 36515
Oy 79 CysThrPheArgValCysPhe-----GluGlnCysCysProTyr---91
Db 36514 CTGCTCTTGCATTTGGATTCAAGACAAATCTTAAGACAGCATCTGCCCTCGCCT 36455
Oy 92 -----ThrPheMetValLysLeuIleAsnGlnAsnCysAspSerAlaArgThrSerAsp 109
Db 36454 ACTGACACCTTCAGAGACAGCCCGCTTAACACAGAGTACAAAGCTCAGCCAGCGAT 36395
Oy 110 AspArgLeuCys 113
Db 36394 ATGAGGCTCTGT 36383

RESULT 13
AE005499 10197 bp DNA linear BCT 21-MAR-2001
LOCUS Escherichia coli O157:H7 EDL933 genome, config 3 of 3, section 118
DEFINITION of 290.
ACCESSION AE005499 AE005174
VERSION AE005499.1 GI:12517151
KEYWORDS
SOURCE
ORGANISM
Escherichia coli O157:H7 EDL933.
Escherichia coli O157:H7 EDL933
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
1 (bases 1 to 10197)
Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D.,
Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,
Postel,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,
Grodeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamoukis,K.,
Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
Welch,R.A. and Blattner,F.R.
Genome sequence of enterohemorrhagic Escherichia coli O157:H7
JOURNAL Nature 409 (6819), 529-533 (2001)
MEDLINE 21074935
PUBMED 11206551
REFERENCE 2 (bases 1 to 10197)
Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D.,
Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,
Postel,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,
Grodeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamoukis,K.,
Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
Welch,R.A., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,

```

TITLE
JOURNAL
FEATURES
SOURCE

Grodeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamoukis,K.,
Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
Welch,R.A. and Blattner,F.R.
Direct Submission
Submitted (22-OCT-2000) Laboratory of Genetics, University of
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
Location/Qualifiers
1. 10197
/organism="Escherichia coli O157:H7 EDL933"
/strain="EDL933"
/serotype="O157:H7"
/db_xref="taxon:155864"
/note="enterohemorrhagic"
107..796
/gene="srlr"
/note="24014"
107..796
/gene="srlr"
/function="regulator: Degradation of small molecules:
Carbon compounds"
/note="Residues 10 to 229 of 229 are 98.63 pct identical
to residues 38 to 257 of 257 from Escherichia coli K-12
Strain MG1655: B2707"
/codon_start=1
/transl_table=11
/product="regulator for gut (srl), glucitol operon"
/protein_id="AAG57814.1"
/db_xref="GI:12517152"
/translation="MTPQALFARKNILEHAGTVIRTYGVVLRKEESPDPIDKTL
INTKKELILEAVGVFIHDSIILIDGSLVLMVPLSRNNITVNTSLVNAIS
LENEQTLIMPGETFKRSAPFGIQAENAFEPHTPEKMGDGLDINGVTYENV
ITVSKAMCARREVIIMADSKRGKSPNVCSLESYDKLITPADIDPAPROALENG
-IDYITIGENSE"
828..1734
/gene="gutQ"
/note="24015"
828..1754
/gene="gutQ"
/function="orf: Unknown function"
/note="Residues 1 to 308 of 308 are 100.00 pct identical
to residues 1 to 308 of 308 from Escherichia coli K-12
Strain MG1655: B2708"
/codon_start=1
/transl_table=11
/product="orf: hypothetical protein"
/protein_id="AAG57815.1"
/db_xref="GI:12517153"
/translation="MLEQESRLPERLGDPEYRANIIHCERGVVSGIGKSGHIG
KRIATLASTGTPAFVHPAEALHGLGMESESDVMEFISGCAKELDIIPRLDK
SIALLMTGKPTSPGLAKAVDIVVERACPMHAPSTSYNTLMGDALAMAYWQ
ARGENEDPARSRPAGALARKLRYHHHMRDDAIPQVALTSVDAKAEISRTGIG
LVAVCAQAOQVGSFTDGDRLRMVLVGGLATTPVNAEMTVGGTTIQSOSRAIDAKETL
MKRTIAPAVVDENGKLTGAINLDPEYAGIIT"
complement(1751..3340)
/gene="ygaA"
/note="24017"
complement(1751..3340)
/gene="ygaA"
/function="putative regulator: Not classified"
/note="Residues 1 to 529 of 529 are 99.43 pct identical to
residues 1 to 529 of 529 from Escherichia coli K-12 Strain
MG1655: B2709"
/codon_start=1
/transl_table=11
/product="putative 2-component transcriptional regulator"
/protein_id="AAG57816.1"
/db_xref="GI:12517154"
/translation="MSSNDYOSNGRYNDIYSQIDSEAKMSFVYIATIEILOGRI
GHODRPRILITLROYLEDASALRLYSDROTPIAIDSLADYIGRRRLALGHPRE
ATARAGDYRFPADSLEPDYDGLIPGOSLKVHACVGLPELFGONLIGALTLDDMP
DDPFDVSEDELILIALAGASMLLIEQLSSQNLPGDAAPFAEVKOTQWIGISPG
MQLKEIIVASIDNLVLISETGKELVAKAHASPRAVNPLVYINCAALPESY
ASELFGHVKAFTGAISNRSGKFEWADNGTLFLEIGELSLALQAKLRLVQYDIO

gene
CDS

RVDGDSRLRVYVLAATNRDLREVLGRFADLPRLSVPLSPVLPBREGDDVTL
 LAGYFCEQCFRLROGLSKRVVLSAGARNLQHSFPGNRELEHAIHRAVYLAIRATRS
 EVLEAQNHPFPEVTPLEPPVAAVPVVKQNLREATEAFORETIROLAOLNHHWAA
 RMLTEVYANLHRLAKRLGLKD
 3452. .4687
 /gene="24018"
 /note="24018"
 3452. .4687
 /gene="24018"
 /function="putative enzyme; Not classified"
 /note="Residues 1 to 285 of 411 are 100.00 pct identical
 to residues 1 to 285 of 479 from Escherichia coli K-12
 strain MG1655: B2710"
 /codon_start=1
 /transl_table=11
 /product="putative flavodoxin"
 /protein_id="AAG57817.1"
 /db_xref="GI:12517155"
 /translation="MSIVVKNNIHWGVRDVEFHFGEYKTLRGSVSYLIREK
 NVLIDVDHKSREFFQNLREIDLADIDYIVNNAEDHAGALTELMQIPTPTIC
 TANAIDISNGHHHPERNVYKGTCTLDIGNKQILFETPLHWPDSMTYLTGDA
 VLFSDAFQCHCDEHLENDEVDQTELEFCQCYRANILTFPSRLVTPKTEILGNL
 PVDMLATSHGVWRDNPQIVELYLMKADYQEDRITIFEDTWSNTPRMADAIAQI
 APTDPRVAVDRSLTRLDAGFEMSLKAKMRPDODALECRESGREIAROMALPLP
 OSTVNTVKEETSAITTDLPGRMCSVCQMIYDPAKGPMDVAPGTSPSEVPDNL
 CECSLGRVPELASEAK"
 4684. .5817
 /gene="y9bd"
 /note="24019"
 4684. .5817
 /gene="y9bd"
 /function="putative enzyme; Not classified"
 /note="Residues 1 to 377 of 377 are 97.87 pct identical to
 residues 1 to 377 of 377 from Escherichia coli K-12 Strain
 MG1655: B2711"
 /codon_start=1
 /transl_table=11
 /product="putative oxidoreductase"
 /protein_id="AAG57818.1"
 /db_xref="GI:12517156"
 /translation="MSNGIIVIGSGFAARQVLKNIKODASIPILTIADSNDEYK
 DLSHVISGQADDLRQTAGFARDOENLTPPTWYDIDAEARVKSNNQWYDK
 LVLATGSAFVPVPGREMLTINSQOEERAEQTOIRARVLYGGGLISBLAND
 CRAGKMWLIDNAAISLAMPYESSRQHLTEGVHLIKLSQLEKIDSGILA
 TLDRSLREVDVAVIATGLRPELTALARRAGLITNGCVDSINADIVLGDCAE
 INGQVLPLOPQLSAMVLAKNLGNNTPLKPAIMVKTPELPHLAGLEPODLR
 WQINTEROGNVAAGVDADQDLRAFYVSEDRMKEARGLKTLVY"
 complement(5945. .8197)
 /gene="hypr"
 /note="24020"
 complement(5945. .8197)
 /gene="hypr"
 /function="regulator; Energy metabolism, carbon: Anaerobic
 respiration"
 /note="Residues 1 to 750 of 750 are 97.86 pct identical to
 residues 1 to 750 of 750 from Escherichia coli K-12 strain
 MG1655: B2712"
 /codon_start=1
 /transl_table=11
 /product="transcriptional regulatory protein"
 /protein_id="AAG57819.1"
 /db_xref="GI:12517157"
 /translation="MAKNTSCGVOLIRGKVGQVGFPEVNOALQNLHGVCNDG
 GVEVRLDEPTFLVOLHQCPLRLAIDSVEREPIQSLPTEFIRSGAGAMTOI
 VPDALCPACLAEMNTPEGRERYRPFICNTGCGPFTIIRAPYRPTVAPAFIC
 ACDEYRDLRFRHAPVACRCEGPEYLEWVSHGHAQOBAIAQAIQIKGNIVAI
 KGIGFHLACDARSNAVATLARKHRAKALAYLVPADGIPDAROLDITPPAIV
 LVDKKYPBELCDIAPGLNEVGMPLPAPLQHLLEQQLCVMTSGNLSGPPAIV
 EQALEDDQGLKDFLHNRIYVORDSDSVRESEMLRSRYVDAIALSPGFENVP
 PVLLGADLKNTPCLVIRGEVYLSOHLGDLSDGQIOTQMRALRLQNIYNTPOYV
 HDHAPGVYSCASEMNLPTQTVLHHNHAACLAEHQWMLDGGVGLITLTDGCMG
 NGALMGECLELVNRECEHGLPAVALHPGDIAAKQWRMLNGLCIEFVEMQYFE
 TASVQONMYVLAIRGINAPLASSGRLFDVAVALGAPATLSIEGEACALAE
 LAASCDEVTHPVPRVNDLDALEFWQWMLNWOAPVQNRMAHFDAIQAQFAALMR

QATMRGTTTVFSGGIHNLRLARLAHLADEFTLLFPOSIPAGDGLSIGGVIAAA
 RMLAGEYONG
 complement(8350. .8877)
 /gene="hyvN"
 /note="24021"
 complement(8350. .8877)
 /gene="hyvN"
 /function="enzyme; Energy metabolism, carbon: Anaerobic
 respiration"
 /note="Residues 1 to 175 of 175 are 99.42 pct identical to
 residues 1 to 175 of 175 from Escherichia coli K-12 strain
 MG1655: B2713"
 /codon_start=1
 /transl_table=11
 /product="involved in electron transport from formate to
 hydrogen, Fe-S centers"
 /protein_id="AAG57820.1"
 /db_xref="GI:12517158"
 /translation="MNRFTIADASKICIGRTCEVACVSHOENODCASLTPETPRI
 HVKGVNISTATVRCQEDAPCANVPGNISDKGVHMERCTGCTVYACPYG
 AMEVVVRPIYRNSGALNVRADAEANKKDCIHRREBGRACMAACPTHALICVDKRL
 EQLSAKRRRTALMP"
 complement(9026. .10039)
 /gene="ascG"
 /note="24022"
 complement(9026. .10039)
 /gene="ascG"
 /function="regulator; Degradation of small molecules:
 Carbon compounds"

Alignment Scores:
 Pred. No.: 33.9 Length: 10197
 Score: 89.50 Matches: 31
 Percent Similarity: 35.098 Conservative: 9
 Best Local Similarity: 27.198 Mismatches: 39
 Query Match: 13.408 Indels: 35
 DB: 1 Gaps: 6

US-09-092-297-17 (1-117) x AE005499 (1-10197)

QY 1 ProleuglnProPro-----ArgAlaMetaProArglyCysile 14
 DB 8319 CCYTTGATGCCAAGAGTCTGTAGATTGTTTGAACATACAGCGCGGTTC 8378
 QY 15 Val-----AlaValAlaAlaIlePheCysIleSerArgLeuLeuLeuSerHisGlyAla 32
 DB 8379 TTTCGCGGCTGAGTGTTCAGTTTATTCG---GATCGACACAAATTAGCGCATGGGTCG 8435
 QY 33 ProValAlaProMetThrPro-----Tyr 40
 DB 8436 GGCACCGCCGACACACGCGCGGCGGCTTTCACGATGTTGCACAGCGTCGATTTATTGG 8495
 QY 41 LeuMetLeuCysGlnProHisIysArgCysGlyAspLysPheIysAspProleuglnHis 60
 DB 8496 CTTCGGCTTTGTTCAGCC-----GTACATTCACCCCTGGCGCCG 8534
 QY 61 CysCysTyrAspAspAlaValValProleuAlaIleArgThiGlnThr-----75
 DB 8535 TGTTCGAAATCACCGGACGATCACCATCTTCATCGCGGCATACGGGACCAACAGCG 8594
 QY 76 -----CysGlyAsnCysThrPheArgValCysPheGluGln 87
 DB 8595 AGTTTTCGACCAATGCAACGTTCTGTCATCATGAACAA 8636
 RESULT 14
 LOCUS AP002562 270365 bp DNA linear BCT 07-MAR-2001
 DEFINITION Escherichia coli O157:H7 DNA, complete genome, section 13/20.
 ACCESSION AP002562 BA000007
 VERSION AP002562.1 GI:13362858
 KEYWORDS Escherichia coli O157:H7 (strain:O157:H7, sub_strain:RIMD 0509952)
 SOURCE DNA.

ORGANISM Escherichia coli O157:H7
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.

REFERENCE
AUTHORS Makino, K., Yokoyama, K., Kubota, Y., Yutsudo, C. H., Kimura, S.,
1 (sites)
Kurokawa, K., Ishii, K., Hattori, M., Tatsuno, I., Abe, H., Iida, T.,
Yamamoto, K., Ohnishi, M., Hayashi, T., Yasunaga, T., Honda, T.,
Sasakawa, C., and Shinagawa, H.
Complete nucleotide sequence of the prophage VP2-Sakai carrying the
verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7
derived from the Sakai outbreak
Genes Genet. Syst. 74 (5), 227-239 (1999)

JOURNAL
MEDLINE
REFERENCE
AUTHORS 2 (sites)
Ohnishi, M., Murata, T., Nakayama, K., Kuhara, S., Hattori, M.,
Kurokawa, K., Yasunaga, T., Yokoyama, K., Makino, K., Shinagawa, H. and
Hayashi, T.
Comparative analysis of the whole set of rRNA operons between an
enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an
Escherichia coli K-12 strain MG1655
Syst. Appl. Microbiol. 23 (3), 315-324 (2000)

JOURNAL
MEDLINE
REFERENCE
AUTHORS 3 (sites)
Yokoyama, K., Makino, K., Kubota, Y., Watanabe, M., Kimura, S.,
Yutsudo, C. H., Kurokawa, K., Ishii, K., Hattori, M., Abe, H., Iida, T.,
Yamamoto, K., Hayashi, T., Yasunaga, T., Honda, T., Sasakawa, C. and
Shinagawa, H.
Complete nucleotide sequence of the prophage VP1-Sakai carrying the
Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli
O157:H7 strain derived from the Sakai outbreak
Gene 258 (1-2), 127-139 (2000)

JOURNAL
MEDLINE
REFERENCE
AUTHORS 4 (sites)
Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K.,
Yokoyama, K., Han, C.-G., Ohtsubo, E., Nakayama, K., Murata, T.,
Tanaka, M., Toke, T., Iida, T., Takami, H., Honda, T., Sasakawa, C.,
Ogasawara, N., Yasunaga, T., Kuhara, S., Shiba, T., Hattori, M. and
Shinagawa, H.
Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12
DNA Res. 8 (1), 11-22 (2001)

JOURNAL
MEDLINE
REFERENCE
AUTHORS 5 (bases 1 to 270365)
Ohnishi, M., Kurokawa, K., Makino, K., Yasunaga, T., Shinagawa, H. and
Hayashi, T.
Direct Submission
Submitted (26-JUN-2000) Ken Kurokawa, Osaka University, Genome
Information Research Center, 3-1, Yamadaoka, Suita, Osaka 565-0871,
Japan (E-mail: kene@gen-info.osaka-u.ac.jp/, Tel: 81-6-6879-8365,
URL: http://www.gen-info.osaka-u.ac.jp/, Fax: 81-6-6879-2047)
genome project.

COMMENT
FEATURES
source
Location/Qualifiers
1..270365
/organism="Escherichia coli O157:H7"
/strain="O157:H7"
/sub_strain="RIMD 0509952"
/db_xref="taxon:83334"
complement(435..1220)
/gene="ECS3388"
complement(435..1220)
/gene="ECS3388"
/note="similar to SEB_ECOLI g1178871 percent identity
98 in 261 aa (Conserved in E.coli K-12)"
/codon_start=1
/transl_table=1
/evidence="not experimental"
/product="enhanced serine sensitivity"
/protein_id="BAB36811.1"
/db_xref="GI:13362859"
/translation="MVLSEKNELEDLLEKAAEPHRAPEFTLESTVWPGTA
OGAEVDSALDLOHWEKEDGTSVPPSTLEAQQAVDEBQAFVMPVPTLEMTIG
ETLELNKLTGKFEFMRISLIGEGNPLSOSLEGGESLILSEVAPPMIDS

gene
complement(1354..2724)
/gene="ECS389"
complement(1354..2724)
/note="similar to PEPB_ECOLI g1178872 percent identity
99 in 456 aa (Conserved in E.coli K-12)"
/codon_start=1
/transl_table=1
/evidence="not experimental"
/product="putative peptidase"
/protein_id="BAB36812.1"
/db_xref="GI:13362860"
/translation="MPPSCNANHLLEKSKCOLAALLORRTKTEAMKITLSTOPAD
AWGSEKATYSINNDGTTLHNGADDLGLQRAARKIDGIGHVQLSGGMPADRCWA
FWQYKAPKGTREKVPDLDADROELDRMLIMIDVPTINAPAEELGPSOLAQWA
DLISVNASDPRVYRTTKGDLREGYMGILHTGSGSERSPVLAIDNPYTGKEAPVY
ACLVGKITEFDGSGYSIKOTAFNDGMSMDGAAVYTGALAFARGLKPKYPLIC
ADNLISGNAFKIGDITITRYNKGKVEYNNNTDAGRLVLDGLIDASQKPEMTIDATL
TGAATLNGNDTHALFSDDLAAGRLASASQENEPFWRPLAEFRHSQLPSPFELN
NGSAFYPGASTAGFLSHFEVNYQGWMLIDCSATYRKAEVQMSAGATGLGYRTI
ANLTA"
complement(2815..3015)
/gene="ECS3390"
complement(2815..3015)
/gene="ECS3390"
/note="identical to YFHJ_ECOLI g1178873 (Conserved in
E.coli K-12)"
/codon_start=1
/transl_table=1
/evidence="not experimental"
/product="hypothetical protein"
/protein_id="BAB36813.1"
/db_xref="GI:13362861"
/translation="WGLKWTDSREIGALYDAYPDLPKTVFTDMQWICLEDIED
DPOASNEKILAILVWIDEAE"
complement(3027..3362)
/gene="ECS3391"
complement(3027..3362)
/note="similar to FER_ECOLI g1178875 percent identity
100 in 111 aa (Conserved in E.coli K-12)"
/codon_start=1
/transl_table=1
/evidence="not experimental"
/product="12Fe-251 ferredoxin"
/protein_id="BAB36814.1"
/db_xref="GI:13362862"
/translation="MPKIVILPHQDLCPDGAVLEANSGETILDAALRNGIEHNACEK
SCACTTCHCIYREGFDSLPESSSEDDMLAKWGLEPSRLSCARVDEDLVVEIPR
XTINHAREH"
complement(3364..5214)
/gene="ECS3392"
complement(3364..5214)
/gene="ECS3392"
/note="similar to HSCA_ECOLI g1178875 percent identity
100 in 616 aa (Conserved in E.coli K-12)"
/codon_start=1
/transl_table=1
/evidence="not experimental"
/product="molecular chaperone"
/protein_id="BAB36815.1"
/db_xref="GI:13362863"
/translation="MALQISEPGLISAAPHQRRLAAGIDITGNSLVATVRSQAEFL
ADHEGRLLPSVYVYOOQSGVGDARTNALDPAITSSYKRLMGLSADIOORYPH
LPYQFASENGLEPIETAGLNVRSADILKARATRETAGERDGVITYPAYF
DLAOROGTKAARLAGLHYRLNLEPPAAIAYIGDSGSCVAVYVLDGGGTISTIL
RLSGVFEVATGDSALGGEDPHLDADYIRBOAGIPDSSDNVQRELLDAANAAT
ALSDADSVTVNVAWQGEISREDFNELIADLYKRTLLACRRALLDACVDEDEVLEVY
VGSSTRVPLVREVERGEFFRPPLTSDIPDKVVAIGAIQADIVGNKPDSEMLLDVY
PLSITGTEMGLEKVEIPRNTTIPVAADPFTFKDGTQMSIHVMOGEBELVDGCS
LARPALRGIDALPAGNAHIVTQYVDADGLLSVTAMKSKSYGVENSIOYKSYGLTDS

41 leuagcgcucysgimprona15ysatcysg1yaspysbheuygrasproleaglnhls 60
 ||| |||||
 Db 187067 ctttcggctttgtcagccc-----gtacatttcagcccctgcgccccc 187105

Db 206609 -----||||| |||::: :: |||
-----CTTGTGCGCTGTGGAAGTGTCCCTTGGCCCTGTGCTGTGCTTGACTTAA 206562

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OW protein - nucleic search, using frame_plus.p2n model

Run on: November 3, 2002, 01:08:01 ; Search time 210 Seconds

(without alignments)
956.567 Million cell updates/sec

Title: US-09-092-297-17
Perfect score: 668
Sequence: 1 PUOPRRAMPRCIVAVFAI.....INONCDASRTSDRLCRSVS 117

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+pn.model -DEV=xlh
-O=/cgn2_1/USPTO/US09092297/runat_29102002_093920_21656/app.query.fasta_1.263
-DB=N-Geneseq_032802 -OFMT=fastap -SUFFIX=32n.rng -MINMATCH=0.1 -LOOPT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcr -NORM=ext HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09092297_@cgn_1_1_0.gruntat_29102002_093920_21656 -NCP=6 -ICPU=3
-NO_XLPHY -NO_WMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONDILOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N-Geneseq_032802:.*
1: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	668	100.0	762	20	AA224391 Human bladder tumo
2	668	100.0	763	20	AA224391 Human bladder tumo
3	668	100.0	763	20	AA224391 Human bladder tumo
4	668	100.0	771	21	AA377668 Human PRO1274 cDNA
5	668	100.0	771	21	AA377668 Human PRO1274 cDNA
6	668	100.0	771	21	AA377668 Human PRO1274 cDNA
7	668	100.0	771	21	AA377668 Human PRO1274 cDNA
8	668	100.0	771	21	AA377668 Human PRO1274 cDNA
9	668	100.0	771	21	AA377668 Human PRO1274 cDNA
10	668	100.0	771	21	AA377668 Human PRO1274 cDNA
11	668	100.0	771	21	AA377668 Human PRO1274 cDNA
12	668	100.0	771	21	AA377668 Human PRO1274 cDNA
13	668	100.0	771	21	AA377668 Human PRO1274 cDNA
14	668	100.0	771	21	AA377668 Human PRO1274 cDNA
15	668	100.0	771	21	AA377668 Human PRO1274 cDNA
16	668	100.0	771	21	AA377668 Human PRO1274 cDNA
17	668	100.0	771	21	AA377668 Human PRO1274 cDNA
18	668	100.0	771	21	AA377668 Human PRO1274 cDNA
19	668	100.0	771	21	AA377668 Human PRO1274 cDNA
20	668	100.0	771	21	AA377668 Human PRO1274 cDNA
21	668	100.0	771	21	AA377668 Human PRO1274 cDNA
22	668	100.0	771	21	AA377668 Human PRO1274 cDNA
23	668	100.0	771	21	AA377668 Human PRO1274 cDNA
24	668	100.0	771	21	AA377668 Human PRO1274 cDNA
25	668	100.0	771	21	AA377668 Human PRO1274 cDNA
26	668	100.0	771	21	AA377668 Human PRO1274 cDNA
27	668	100.0	771	21	AA377668 Human PRO1274 cDNA
28	668	100.0	771	21	AA377668 Human PRO1274 cDNA
29	668	100.0	771	21	AA377668 Human PRO1274 cDNA
30	668	100.0	771	21	AA377668 Human PRO1274 cDNA
31	668	100.0	771	21	AA377668 Human PRO1274 cDNA
32	668	100.0	771	21	AA377668 Human PRO1274 cDNA
33	668	100.0	771	21	AA377668 Human PRO1274 cDNA
34	668	100.0	771	21	AA377668 Human PRO1274 cDNA
35	668	100.0	771	21	AA377668 Human PRO1274 cDNA
36	668	100.0	771	21	AA377668 Human PRO1274 cDNA
37	668	100.0	771	21	AA377668 Human PRO1274 cDNA
38	668	100.0	771	21	AA377668 Human PRO1274 cDNA
39	668	100.0	771	21	AA377668 Human PRO1274 cDNA
40	668	100.0	771	21	AA377668 Human PRO1274 cDNA
41	668	100.0	771	21	AA377668 Human PRO1274 cDNA
42	668	100.0	771	21	AA377668 Human PRO1274 cDNA
43	668	100.0	771	21	AA377668 Human PRO1274 cDNA
44	668	100.0	771	21	AA377668 Human PRO1274 cDNA
45	668	100.0	771	21	AA377668 Human PRO1274 cDNA

ALIGNMENTS

RESULT 1
AA224391 standard; cDNA; 762 BP.
AA224391: 14-FEB-2000 (first entry)
Human bladder tumour cDNA library derived EST 3.
Expressed sequence tag; human; bladder; tumour; cancer; cytostatic;
treatment; gene therapy; EST; ss.
Homo sapiens.
DE1981619-A1.
28-OCT-1999.

[illegible]

```

ID      AAX01686 standard; DNA; 763 BP.
XX
XX AC      AAX01686;
XX DT      21-APR-1999 (first entry)
XX DE      Urinary tract tissue library BL172 gene full length clone 1548381H.
XX DE      BL172; urinary tract; marker; cancer; recombinant; human; ss.
XX OS      Homo sapiens.
XX PN      M09855656-A1.
XX PD      10-DEC-1998.
XX PF      05-JUN-1998; 98WO-US11693.
XX PR      05-JUN-1997; 97US-0869579.
XX PA      (ABBO ) ABBORT LAB.
XX PI      Billing-medel PA, Cohen M, Colpitts TL, Friedman PN;
XX PI      Gordon J, Granadosen, Hodges SC, Klass MR, Kratochvil JD;
XX PI      Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;
XX PI      WPI, 1999-045802/04.
XX
XX PT      New purified polynucleotide BL172 derivatives and encoded
XX PT      polypeptides - useful in the identification of markers, which are
XX PT      indicative of urinary tract diseases or conditions
XX PS      Claim 1; Page 89-90; 114pp; English.
XX
XX      This represents a full-length sequence of a clone specific for urinary
XX      tract tissue library BL172. The BL172 polynucleotides (AAX01683-87) are
XX      used to detect target BL172 polynucleotides and BL172 mRNA and the BL172
XX      polypeptides are used to detect BL172 antigens/antibodies in a test
XX      sample. The identification of certain markers in these methods are
XX      indicative of the presence of urinary tract disease, especially cancer.
XX      Additionally, the polypeptides are used to detect antibodies that bind
XX      specifically to a BL172 epitope derived from a BL172 polypeptide. Host
XX      cells containing a recombinant expression vector comprising the BL172
XX      polynucleotide sequences are used to produce the polypeptides containing
XX      BL172 epitopes. The methods aid in diagnosis, detection, staging,
XX      monitoring, prognosis, in vivo imaging, prevention and treatment
XX      of diseases or conditions associated with BL172, especially urinary
XX      tract cancer.
XX
XX      Sequence 763 BP; 162 A; 231 C; 195 G; 175 T; 0 other;
XX
XX
XX      Alignment Scores:
XX      Pred. NO.:
XX      Score: 1,5e-64 Length: 763
XX      Percent Similarity: 668.00 Matches: 117
XX      Percent Similarity: 100.008 Conservative: 0
XX      Best Local Similarity: 100.008 Mismatches: 0
XX      Query Match: 100.00% Indels: 0
XX      DB: 20 Gaps: 0
XX
XX
XX      US-09-092-297-17 (1-117) x AAX01686 (1-763)
XX
XX      QY      1 ProlenglnProProlargalametAlaProargGlyCysIleValAlaValPhealaIle 20
XX      Db      3 CCACTGGACACCACCAAGGCATGGCCCGGAGGCGCATGCTAGTGTCTTGGCCATT 62
XX      QY      21 PheCysIleSerArgIleuLeuCySerHisGlyAlaProValAlaProMetThrProTyr 40
XX      Db      63 TTTCGATCTTCACAGCTCTCTGCTCAAGAGGAGCCCACTGAGTCCCTTAC 122
XX      QY      41 LeuMetLeuCysGlnProHisLysArgCysGlyAspLysPheThrYrAspProLeuGlnHis 60
XX      Db      123 CTGATGCTGTGGCCAGCCACACAAAGATGTGGGGAGCAAGTTTCACAGACCCCTGCAGAC 182

```


PR 29-OCT-1999; 99US-0162506.
 XX (GETH) GENENTECH INC.
 PA Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ, Goddard A;
 PI Godowski PJ, Gurney AL, Klein RD, Kuo SS, Paoni NF, Smith V;
 PI Matanabe CK, Williams PM, Wood WI.
 XX WPI: 2000-412154/35.
 DR P-PSDB; AAB24425.
 XX
 PT Nucleic acids encoding PRO polypeptides useful for preventing,
 PT diagnosing and treating a cardiovascular, endothelial or
 XX angiotensin disorders in mammals -
 PS Claim 61, Fig 75; 315pp; English.
 CC The present invention describes nucleic acids encoding PRO polypeptides
 CC useful for preventing, diagnosing and treating a cardiovascular, endothelial or
 CC modulating cell proliferation, angiogenesis and cardiovascularisation,
 CC and for identifying agonists and antagonists of these processes. The
 CC nucleic acids and the proteins they encode may be used in the
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate PRO expression such as cardiovascular, endothelial or
 CC angiogenic disorders in mammals (e.g. atherosclerosis, cancers and
 CC cardiac hypertrophy). For example, the nucleic acids (NAs) and vectors
 CC containing them and the PRO polypeptide may be used to treat disorders
 CC associated with decreased PRO expression. AAB24435 to AAB24438 to AAB24435 represent nucleotide and protein sequences used in
 CC the exemplification of the present invention.
 XX
 SQ Sequence 771 BP; 169 A; 231 C; 195 G; 176 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.52e-64 Length: 771
 Score: 668.00 Matches: 117
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-092-297-17 (1-117) x AAA77668 (1-771)
 QY 1 ProLeuGlnProPheArgAlaMetAlaProArgGlyCysIleValAlaValPheAlaIle 20
 DB 3 CCAGTGCACACACACAGAGCCATGCTCCCGAGGCTGCATGCTGCTTTGCCATT 62
 QY 21 PheCysIleSerArgLeuLeuCysSerHisGlyAlaProValAlaProMetThrProTyr 40
 DB 63 TTCTGCATCTCCAGAGCTCCCTGCTCACAGAGGCCCGCCAGTGGCCCAAGTCTTAC 122
 QY 41 LeuMetLeuGlnProHisIleArgCysGlyAspLysPheTyrAspProLeuGlnHis 60
 DB 123 CTATGCTGTCGACGACACAGAGAGTGTGGGACAAAGTCTACACACCCCGCAGAC 182
 QY 61 CysCysTyrAspAspAlaValAlaProLeuAlaArgThrGlnThrCysGlyAsnCysThr 80
 DB 183 TGTTCGTATGATGATCCCTGCGCCCTTGGCCAGACACGCTGTGAACCTGCAC 242
 QY 81 PheArgValCysPheGlnGlnCysCysProTyrPheMetValLysLeuIleAsnGln 100
 DB 243 TTCACAGTCTGCTTGAGAGTGCCTCCCTGACCTTCACTGATGAACCTGATAACAG 302
 QY 101 AsnCysAspSerAlaArgThrSerAspAspArgLeuCysArgSerValSer 117
 DB 303 AACTCGCGCTACGCCGAGACCTCGATGACAGCGCTTGTCCAGAGTGCAGC 353
 RESULT 5
 ID AAA37062 standard; cDNA; 771 BP.
 XX AAA37062;
 AC AAA37062;

XX 08-AUG-2000 (first entry)
 DE Human PRO1274 (UNC644) cDNA sequence SEQ ID NO:137.
 XX
 DE Human: PRO polypeptide; membrane bound protein; receptor; diagnosis;
 KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening;
 KW ss.
 XX Homo sapiens.
 XX WO200012708-A2.
 PD 09-MAR-2000.
 XX
 PF 01-SEP-1999; 99WO-US20111.
 XX 01-SEP-1998; 98US-0098716.
 PR 01-SEP-1998; 98US-0098749.
 PR 01-SEP-1998; 98US-0098750.
 PR 02-SEP-1998; 98US-0098803.
 PR 02-SEP-1998; 98US-0098821.
 PR 02-SEP-1998; 98US-0098843.
 PR 09-SEP-1998; 98US-0099536.
 PR 09-SEP-1998; 98US-0099596.
 PR 09-SEP-1998; 98US-0099598.
 PR 09-SEP-1998; 98US-0099602.
 PR 10-SEP-1998; 98US-0099642.
 PR 10-SEP-1998; 98US-0099741.
 PR 10-SEP-1998; 98US-0099754.
 PR 10-SEP-1998; 98US-0099792.
 PR 10-SEP-1998; 98US-0099808.
 PR 10-SEP-1998; 98US-0099812.
 PR 10-SEP-1998; 98US-0099815.
 PR 15-SEP-1998; 98US-0099816.
 PR 15-SEP-1998; 98US-0100385.
 PR 15-SEP-1998; 98US-0100388.
 PR 16-SEP-1998; 98US-0100390.
 PR 16-SEP-1998; 98US-0100584.
 PR 16-SEP-1998; 98US-0100627.
 PR 16-SEP-1998; 98US-0100651.
 PR 16-SEP-1998; 98US-0100662.
 PR 17-SEP-1998; 98US-0100664.
 PR 17-SEP-1998; 98US-0100683.
 PR 17-SEP-1998; 98US-0100684.
 PR 17-SEP-1998; 98US-0100710.
 PR 17-SEP-1998; 98US-0100711.
 PR 17-SEP-1998; 98US-0100919.
 PR 18-SEP-1998; 98US-0100930.
 PR 18-SEP-1998; 98US-0100848.
 PR 18-SEP-1998; 98US-0100849.
 PR 18-SEP-1998; 98US-0100849.
 PR 18-SEP-1998; 98US-0101014.
 PR 18-SEP-1998; 98US-0101068.
 PR 22-SEP-1998; 98US-0101071.
 PR 23-SEP-1998; 98US-0101079.
 PR 23-SEP-1998; 98US-0101471.
 PR 23-SEP-1998; 98US-0101472.
 PR 23-SEP-1998; 98US-0101474.
 PR 23-SEP-1998; 98US-0101475.
 PR 23-SEP-1998; 98US-0101476.
 PR 24-SEP-1998; 98US-0101477.
 PR 24-SEP-1998; 98US-0101738.
 PR 24-SEP-1998; 98US-0101741.
 PR 24-SEP-1998; 98US-0101743.
 PR 24-SEP-1998; 98US-0101915.
 PR 29-SEP-1998; 98US-0102207.
 PR 29-SEP-1998; 98US-0102240.
 PR 29-SEP-1998; 98US-0102307.
 PR 29-SEP-1998; 98US-0102330.
 PR 29-SEP-1998; 98US-0102331.

PR 30-SEP-1998; 98US-0102484.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102685.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103314.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103395.
PR 07-OCT-1998; 98US-0103396.
PR 07-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103633.
PR 08-OCT-1998; 98US-0103678.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103711.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.
PR 22-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105693.
PR 26-OCT-1998; 98US-0105694.
PR 27-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105881.
PR 27-OCT-1998; 98US-0105882.
PR 28-OCT-1998; 98US-0106062.
PR 28-OCT-1998; 98US-0106023.
PR 28-OCT-1998; 98US-0106029.
PR 28-OCT-1998; 98US-0106030.
PR 28-OCT-1998; 98US-0106033.
PR 28-OCT-1998; 98US-0106178.
PR 29-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106384.
PR 29-OCT-1998; 98US-0108500.
PR 30-OCT-1998; 98US-0106464.
PR 03-NOV-1998; 98US-0106856.
PR 03-NOV-1998; 98US-0106902.
PR 03-NOV-1998; 98US-0106905.
PR 03-NOV-1998; 98US-0106919.
PR 03-NOV-1998; 98US-0106932.
PR 03-NOV-1998; 98US-0106934.
PR 10-NOV-1998; 98US-0107783.
PR 17-NOV-1998; 98US-0108775.
PR 17-NOV-1998; 98US-0108779.
PR 17-NOV-1998; 98US-0108787.
PR 17-NOV-1998; 98US-0108788.
PR 17-NOV-1998; 98US-0108801.
PR 17-NOV-1998; 98US-0108802.
PR 17-NOV-1998; 98US-0108806.
PR 17-NOV-1998; 98US-0108807.
PR 17-NOV-1998; 98US-0108867.
PR 17-NOV-1998; 98US-0108925.
PR 18-NOV-1998; 98US-0108848.
PR 18-NOV-1998; 98US-0108849.
PR 18-NOV-1998; 98US-0108850.
PR 18-NOV-1998; 98US-0108851.
PR 18-NOV-1998; 98US-0108852.
PR 18-NOV-1998; 98US-0108858.
PR 18-NOV-1998; 98US-0108904.
XX (GETH) GENENTECH INC.
XX Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI.
XX MPI: 2000-237871/20.
XX DR P-PSDB; AA199380.

XX New mammalian DNA sequences encoding transmembrane, receptor or
PT secreted PRO polypeptides, useful for screening of potential peptide or
PR small molecule inhibitors of the relevant receptor/ligand interactions
XX
XX Claim 2; Fig 81; 773pp; English.
XX
CC AAA37022 to AAA37144 encode the new isolated human transmembrane,
CC receptor or secreted PRO polypeptides given in AA199340 to AA199462. The
CC transmembrane and receptor PRO proteins can be used for screening of
CC potential peptide or small molecule inhibitors of the relevant
CC receptor/ligand interactions. The polypeptides and nucleotide sequences
CC encoding them have various industrial applications, including uses as
CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
CC PCR primers and hybridisation probes used in the isolation of the PRO
CC polypeptides from the present invention.
XX
SQ Sequence 771 BP; 169 A; 231 C; 195 G; 176 T; 0 other;

Alignment scores:
Pred. No: 1.52e-64 Length: 771
Score: 668.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 21

US-09-092-297-17 (1-117) x AAA37062 (1-771)

QY 1 ProLeuGlnProProArgAlaMetAlaProArgGlyCysIleValAlaValPheAlaIle 20
Db 3 CCACGTGCAACCCACCCAGAGCCATGCTCCCGAGGCTCATCGTAGCTGTTGCCATT 62

QY 21 PheCysIleSerArgLeuLeuLeuCysSerHisGlyAlaProValAlaPromethProTyr 40
Db 3 TTCTGCATCTCCAGGCTCTCTGCTCACAGAGCCGCCAGGCGCCCATACATCTTAC 122

QY 41 LeuMetLeuGlyGlnProHisArgGlyAspIleValPheTyrAspProLeuGlnHis 60
Db 123 CTGATGCTGTCGCCAGCCACACAGATGTTGGGACAAAGTTCTACGACCCCTGACAGC 182

QY 61 CysCysTyrAspAspAlaValAlaValProLeuAlaArgThrGlnThrCysGlyAsnCysThr 80
Db 183 TGTTCGTATGATGATGCGCTGCTGCCAGAGCCAGACGTGGAACTGCAC 242

QY 81 PheArgValCysPheGlnGlnCysCysSerProThrPheMetValIleAsnGln 100
Db 243 TTCAGAGTCTGCTTGACAGCTGCTGCCCTGGACCTTCAGTGAAGCTGATAACCG 302

QY 101 AsnCysAspSerAlaArgThrSerAspAspArgLeuCysArgSerValSer 117
Db 303 AACCTGCGACTCAGCCCGGAGCTCGGATGACAGGCTTGTGCAATGTCAGC 353

RESULT 6
AAS21491
ID AAS21491 standard; cDNA; 771 BP.
XX
AC AAS21491;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human cDNA sequence encoding for PRO1274 polypeptide.
XX
KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW breast; prostate; cervical; tumour necrosis factor- α 1pha; TNF- α 1pha;
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200140466-A2.
XX
PD 07-JUN-2001.

XX 01-DEC-2000; 2000WO-US32678.
 XX 01-DEC-1999; 99WO-US28301.
 PR 01-DEC-1999; 99WO-US28634.
 PR 02-DEC-1999; 99WO-US28551.
 PR 02-DEC-1999; 99WO-US28564.
 PR 02-DEC-1999; 99WO-US28565.
 PR 09-DEC-1999; 99US-0170262.
 PR 16-DEC-1999; 99WO-US30095.
 PR 20-DEC-1999; 99WO-US30911.
 PR 30-DEC-1999; 99WO-US30999.
 PR 06-JAN-2000; 2000WO-US0277.
 PR 06-JAN-2000; 2000WO-US00376.
 PR 11-FEB-2000; 2000WO-US03565.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 22-FEB-2000; 2000WO-US04342.
 PR 24-FEB-2000; 2000WO-US04914.
 PR 24-FEB-2000; 2000WO-US05004.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 20-MAR-2000; 2000WO-US07377.
 PR 21-MAR-2000; 2000WO-US07532.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 10-NOV-2000; 2000WO-US30873.
 XX (GETH) GENENTECH INC.
 XX Baker KP, Beresht M, Deforge L, Desnoyers L, Filvaroff E, Gao W,
 PI Gerltsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 DR MPI: 2001-408281/43.
 DR P-PSDB; AA012419.
 XX Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing
 PT lung polypeptides, and detect the presence of mammalian tumours e.g.
 XX lung, breast, prostate, cervical
 PS Claim 3: Fig 495; 813pp; English.
 CC AAS21244-AAS21518 encode for novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumor necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, or the release of cytokine from peripheral blood
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIa. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 XX
 SQ Sequence 771 BP; 165 A; 231 C; 195 G; 176 T; 0 other;

Alignment Scores:
 Pred. No.:
 Score:

1.52e-64 Length: 771
 668.00 Matches: 117

Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 22 Gaps: 0
 US-09-092-297-17 (1-117) x AAS21491 (1-771)
 QY 1 ProleuglnProporArgAlaMetAlaProArgGlyCysIleValAlaValPheAlaIle 20
 Db 3 CCACTGCACACACCCAGCCGATGCTCCCGAGGCTGCATCGAGCTGTTCGCCATT 62
 QY 21 PheCysIleSerArgLeuLeuCysSerHisGlyAlaProValAlaProMetThrProTyr 40
 Db 63 TTTCGATCTCCAGGCTCCTCTGCTCACAGGAGCCACAGGCCCCCATCTATCTCTTAC 122
 QY 41 LeuMetLeuCysGlnProHisLysArgCysGlyAspLysPheTyrAspProleuglnHis 60
 Db 123 CTGATGCTGTGCCAGCCACACAGAGATGTGGGACAAAGTTCTACGACCCCTGCAGCAC 182
 QY 61 CysCysTyrAspAspAlaValAlaValProleuAlaArgThrGlnThrCysGlyAsnCysThr 80
 Db 183 TGTTCCTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242
 QY 81 PheArgValCysPheGluGlnCysCysProTyrThrPheMetValLysIleAsnGln 100
 Db 243 TTCGAGCTGTGCTTGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
 QY 101 AsnCysAspSerAlaArgThrSerAspAspArgLeuCysArgSerValSer 117
 Db 303 AACTGCACACACCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 353
 RESULT 7
 AAC85962
 ID AAC85962 standard; cDNA: 771 BP.
 AC AAC85962;
 AC AAC85962;
 DT 22-AUG-2001 (first entry)
 DE Native sequence of PRO1274 cDNA, clone DNA64889-1541.
 KW PRO: PRO1081; PRO1274; PRO10272; proliferation; T-lymphocyte;
 KW PRO1199; PRO1556; PRO4401; PRO10268; inhibition; stimulation;
 KW infiltration; mononuclear cell; eosinophil; erythema multiforme;
 KW polymorphonuclear neutrophil; PMN; antibody; immune-related disorder;
 KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
 KW juvenile chronic arthritis; spondyloarthritis; systemic sclerosis;
 KW idiopathic inflammatory myopathy; Sjogren's syndrome; skin disease;
 KW autoimmune thrombocytopenia; autoimmune haemolytic anaemia; asthma;
 KW immune-mediated renal disease; demyelination; diabetes mellitus; allergy;
 KW peripheral nervous system; idiopathic demyelinating polyneuropathy;
 KW Guillain-Barre syndrome; hepatobiliary disease; eosinophilic pneumonia;
 KW chronic active hepatitis; primary biliary cirrhosis; allergic rhinitis;
 KW granulomatous hepatitis; sclerosing cholangitis; food hypersensitivity;
 KW inflammatory bowel disease; gluten-sensitive enteropathy; urticaria;
 KW Whipple's disease; idiopathic pulmonary fibrosis; contact dermatitis;
 KW psoriasis; atopic dermatitis; hypersensitivity pneumonitis;
 KW graft rejection; graft-versus-host disease; ss.
 KW
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FH CDS 24..356
 FT /*tag= a
 FT /product= "PRO1274 polypeptide"
 FT 24..95
 FT /*tag= b
 FT mat_peptide 96..353
 FT /*tag= c
 XX
 XX WO200140465-A2.

PD 07-JUN-2001.
 XX 10-NOV-2000; 2000MO-US30873..
 XX 30-NOV-1999; 99MO-US28313..
 PR 09-DEC-1999; 99US-0170262..
 PR 23-DEC-1999; 99US-0172059..
 PR 11-JAN-2000; 2000US-0175481..
 PR 20-JAN-2000; 2000US-0177118..
 PR 18-FEB-2000; 2000MO-US04342..
 PR 03-MAR-2000; 2000US-0187202..
 PR 30-MAY-2000; 2000MO-US14941..
 PR 05-JUN-2000; 2000US-0209832..
 PR 24-AUG-2000; 2000MO-US23328..
 XX (GENTECH) GENENTECH INC..
 PA Fong S, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
 PI Tunas D, Watanabe CK, Wood WI, Zhang Z;
 XX P-PSDB; AAB47292.
 DR MPI: 2001-381384/40.
 XX P-PSDB; AAB47292.
 PT Isolated PRO polypeptide useful for treat or diagnose an immune-related
 PT disorder e.g. arthritis, asthma, allergy, diabetes or psoriasis -
 XX
 PS Claim 2; Fig 3; 124pp; English.
 XX The sequences given in AAC85961-69 encode PRO polypeptides. PRO1081,
 CC PRO1274 and PRO1072 stimulate the proliferation of T-lymphocytes and
 CC PRO1199, PRO1556, PRO4401 and PRO10268 inhibit the proliferation of
 CC T-lymphocytes. PRO1754 and PRO9912 act to enhance the infiltration of
 CC mononuclear cells, eosinophils or polymorphonuclear neutrophils (PMN)
 CC into the tissue of a mammal. The PRO cDNA's and antibodies which
 CC bind to them, are used to treat an immune-related disorder in a
 CC mammal. Such disorders include systemic lupus erythematosus,
 CC rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis, a
 CC spondyloarthropathy, systemic sclerosis, an idiopathic inflammatory
 CC myopathy, Sjogren's syndrome, systemic vasculitis, sarcoidosis,
 CC autoimmune haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis,
 CC diabetes mellitus, immune-mediated renal disease, a demyelinating
 CC disease of the central or peripheral nervous system, idiopathic
 CC demyelinating polyneuropathy, Guillain-Barre syndrome, a chronic
 CC inflammatory demyelinating polyneuropathy, a hepatobiliary disease,
 CC infectious or autoimmune chronic active hepatitis, primary biliary
 CC cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
 CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
 CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
 CC disease, erythema multiforme, contact dermatitis, psoriasis, an
 CC allergic disease, asthma, allergic rhinitis, atopic dermatitis, food
 CC hypersensitivity, urticaria, an immunologic disease of the lung,
 CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
 CC pneumonitis, a transplantation associated disease, graft rejection or
 CC graft-versus-host disease.
 XX
 SQ Sequence 771 BP; 169 A; 231 C; 195 G; 176 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.52e-64 Length: 771
 Score: 668.00 Matches: 117
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-092-297-17 (1-117) x AAC85962 (1-771)
 QY 1 ProleuglnProProlAvalMetalProArglyCysIleValAlaValPheAlaIle 20
 Db 3 CCAATGCAACACCAAGGCAATGGCTCCCGAGGCTGATGCTGCTTTCATT 62
 QY 21 PheCysIleSerArgLeuLeuCysSerHisGlyAlaProValAlaPromethrProtyr 40

Db 63 TTCGATCTCCAGGCTCCTCTGCTCAGACGAGCCCGAGTGGCCCCCATGATCTCTTAC 122
 QY 41 LeuMetLeuCysGlnProHisIlyArgCysGlyAspIlyAspIlyAspProleuglnHis 60
 Db 123 CTGATGCTGTGTCCACCCACACAGAGATGTGGGACAAATTTCAGACCCCTCGACGAC 182
 QY 61 CysCysTyRAspAspAlaValAlaValProleuAlaArgThnGlnThrCysGlyAsnCysThr 80
 Db 183 TGTGCTATGATGATGATGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242
 QY 81 PheArgValAlCysPheGlnGlnCysCysProTyrThrPheMetValIlyLeuIleAsnGln 100
 Db 243 TTCAGAGCTGCTGCTTTCAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
 QY 101 AsnCysAspSerAlaArgThrSerAspAspArgLeuCysArgSerValSer 117
 Db 303 AACTGCGACTCAGCCGAGCTCGATGACAGGCTTGTGGCAGTGTACG 353
 RESULT 8
 AAF54298
 ID AAF54298 standard; DNA; 771 BP.
 AC AAF54298;
 XX
 XX 02-APR-2001 (first entry)
 DT
 XX DNA encoding protein of the invention #41.
 DE
 XX Secreted; transmembrane; gene therapy; ss.
 KM
 XX Unidentified.
 OS
 XX WO200078961-A1.
 PN
 XX 28-DEC-2000.
 PD
 XX 18-FEB-2000; 2000MO-US04342.
 PF
 XX 23-JUN-1999; 99US-0141037..
 PR 20-JUL-1999; 99US-0144758..
 PR 26-JUL-1999; 99US-0145698..
 PR 01-SEP-1999; 99MO-US20111..
 PR 28-OCT-1999; 99US-0162506..
 PR 30-NOV-1999; 99MO-US28313..
 PR 02-DEC-1999; 99MO-US28551..
 PR 16-DEC-1999; 99MO-US30095..
 PR 05-JAN-2000; 2000MO-US00219..
 PR 06-JAN-2000; 2000MO-US00376..
 XX
 XX (GENTECH) GENENTECH INC..
 PA Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
 PI Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
 PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tunas D;
 PI Watanabe CK, Williams PM, Wood WI;
 XX
 DR MPI: 2001-071395/08.
 DR
 XX Secreted and transmembrane proteins and nucleic acids designated PRO,
 XX useful as hybridization probes, in chromosome and gene mapping and gene
 XX therapy -
 XX
 XX Claim 2; Fig 81; 787pp; English.
 XX The present invention relates to secreted and transmembrane proteins.
 XX These proteins and the DNA encoding them may be used as hybridization
 XX probes, in chromosome and gene mapping and in the generation of
 XX anti-sense RNA and DNA. They may also be used to generate either
 XX transgenic animals or knockout animals which are in turn useful for
 XX development and screening of therapeutically useful reagents.
 XX The nucleic acids may also be used in gene therapy.
 XX
 XX Sequence 771 BP; 169 A; 231 C; 195 G; 176 T; 0 other;

Alignment Scores:

Pred. No.: 1.52e-64
 Score: 668.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 22
 Length: 771
 Matches: 117
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-09-092-297-17 (1-117) x AAF54298 (1-771)

OY 1 PROLEUGINPROBROARGALAMETALAPROARGGLYCYSILLEYALALVALPHEALALLE 20
 DB 3 CCACGTGCAACCCAGAGCCATGGCTCCCGAGGCTGCATCGTAGCTGTTCCTTGCATT 62
 OY 21 PHECYSILLESERARGLEULEUCYSSERHISGLYALAPROVALALAPROMETHPRORY 40
 DB 63 TTCTGCTATCTCAGGCTCTCTGCTACACGAGGCCCAAGTGGCCCATGATGCTCTTAC 122
 OY 41 LEUWELLEUCYSGINPROHISLYSARGCYSGLYASPLYSPHYRYASPPROLEUGINHIS 60
 DB 123 CTGATGCTGTGCCAGCCACACAGATGTGGGACAAGTTCTACGACCCCTGACAC 182
 OY 61 CYSYSTYRASPAPALAVALVALPROLEUALAARGTHRGLNTHRCYSGLYASNYSTR 80
 DB 183 TGTTCATATGATGATGCGCTGCTGCTGCGCCAGGACCCAGACGCTGTGAAACTGCACC 242
 OY 81 PHEARGVALCYSBHEGLUGINCYSYSPROTRPTHRPHMETVALYLSLEULIASGLN 100
 DB 243 TTACAGATCTGCTTTAGAGAGTGTGCTGCCCTGACCTTCAATGCTGAAGCTGATTAAC 302
 OY 101 ASNCYSASPSEALAAARGTHRSEARASAPARGLEUCYSARGSERVALSER 117
 DB 303 AACTGGAGCTCAGCCCGGACCTGGATGACAGGCTTGTGCGAGTGTAC 353

RESULT 9

AAA08111
 ID AAA08111 standard; DNA; 914 BP.

AC AAA08111;

DT 22-JUN-2000 (first entry)

DE Human bladder specific gene 22638 clone 819141H SEQ ID NO:1.

KW Human; bladder specific gene; BSG; bladder cancer; diagnosis;

KW detection; metastasis; cytosolic; ds.

OS Homo sapiens.

PN WO200012761-A1.

PD 09-MAR-2000.

PF 01-SEP-1999; 99WO-US20003.

PR 02-SEP-1998; 98US-0098825.

PA (DIAD-) DIADEXUS LLC.

PI Salceda S. Sun Y, Recipon H, Caferkey R;

DR WPI; 2000-256658/22.

XX Detecting, diagnosing metastasis, staging, monitoring, imaging and

PT treating bladder cancer, involves measuring bladder specific gene (BSG)

XX levels in cells or body fluids

PS Claim 6; Page 25-26; 31pp; English.

XX A method has been developed for detecting, diagnosing metastasis,

CC staging and monitoring the onset of metastasis or a change in stage

CC of bladder cancer, by measuring bladder specific gene (BSG) levels in

CC cells, tissues or body fluids. An increase in BSG levels compared to
 CC normal human control is associated with the presence or progression of
 CC bladder cancer. A decrease is associated with cancer regression. The
 CC method comprising measuring BSG is useful for detecting, diagnosing
 CC metastasis, staging and monitoring the onset of metastasis or a change
 CC in stage of bladder cancer. An antibody against BSG (1) labeled with
 CC paramagnetic ions or an isotope is useful for imaging bladder cancer
 CC and (1), especially conjugated to a cytotoxic agent is useful for
 CC treating bladder cancer. The method is efficient in discriminating
 CC between bladder cancer which has metastasized and the bladder cancer
 CC sequence used in the exemplification of the present invention.

Sequence 914 BP: 193 A; 265 C; 237 G; 218 T; 1 other;

Alignment Scores:

Pred. No.: 1.9e-64
 Score: 668.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 21
 Length: 914
 Matches: 117
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-09-092-297-17 (1-117) x AAA08111 (1-914)

OY 1 PROLEUGINPROBROARGALAMETALAPROARGGLYCYSILLEYALALVALPHEALALLE 20
 DB 35 CCACGTGCAACCCAGAGCCATGGCTCCCGAGGCTGCATCGTAGCTGTTCCTTGCATT 94
 OY 21 PHECYSILLESERARGLEULEUCYSSERHISGLYALAPROVALALAPROMETHPRORY 40
 DB 95 TTCTGCTATCTCAGGCTCTCTGCTACACGAGGCCCAAGTGGCCCATGATGCTCTTAC 154
 OY 41 LEUWELLEUCYSGINPROHISLYSARGCYSGLYASPLYSPHYRYASPPROLEUGINHIS 60
 DB 155 CTGATGCTGTGCCAGCCACACAGATGTGGGACAAGTTCTACGACCCCTGACAC 214
 OY 61 CYSYSTYRASPAPALAVALVALPROLEUALAARGTHRGLNTHRCYSGLYASNYSTR 80
 DB 215 TGTTCATATGATGATGCGCTGCTGCTGCGCCAGGACCCAGACGCTGTGAAACTGCACC 274
 OY 81 PHEARGVALCYSBHEGLUGINCYSYSPROTRPTHRPHMETVALYLSLEULIASGLN 100
 DB 275 TTACAGATCTGCTTTAGAGAGTGTGCTGCCCTGACCTTCAATGCTGAAGCTGATTAAC 334
 OY 101 ASNCYSASPSEALAAARGTHRSEARASAPARGLEUCYSARGSERVALSER 117
 DB 335 AACTGGAGCTCAGCCCGGACCTGGATGACAGGCTTGTGCGAGTGTAC 385

RESULT 10

AA298158
 ID AA298158 standard; cDNA; 748 BP.

AC AA298158;

DT 11-MAY-2000 (first entry)

DE Human signal peptide containing protein HSPF-50 cDNA SEQ ID NO:184.

KW Human; signal peptide-containing protein; HSPF; diagnosis; cancer;

KW inflammation; cardiovascular disease; anticancer; anti-inflammatory;

KW antimicrobial; nootropic; neuroprotective; cardiovascular; hepatocytic;

KW antistimulant; gene therapy; cell proliferation; neurological disorder;

KW reproductive disorder; developmental disorder; arteriosclerosis;

KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anemia;

KW asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;

KW Parkinson's disease; Huntington's disease; ovulatory defect;

XX muscular dystrophy; ss.

OS Homo sapiens.

XX WO200000610-A2.

XX	06-JAN-2000.
PF	25-JUN-1999; 99WO-US144484.
XX	
PR	26-JUN-1998; 98US-0090762.
PR	31-JUL-1998; 98US-0094983.
PR	01-OCT-1998; 98US-0102686.
PR	11-DEC-1998; 98US-0112129.
XX	
PA	(INCY-) INCYTE PHARM INC.
PI	Lai P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR;
PI	Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;
PI	Bardman O;
XX	
DR	WPI: 2000-160673/14.
DR	P-PSDB: AAY87273.
XX	
PT	New human signal peptide-containing proteins useful in treatment,
PT	prevention and diagnosis of e.g. cancer, inflammation and
PT	cardiovascular disease
XX	
PS	Claim 9; Page 281; 327pp: English.
XX	
CC	AAZ98109 to AAZ98242 encode AAY87224 to AAY87357 which represent the
CC	human signal peptide-containing proteins HSPP-1 to HSPP-134. HSPPs have
CC	anticancer, anti-inflammatory, antimicrobial, nootropic, hepatotrophic,
CC	neuroprotective, cardiovascular and antistimatic activities, and can
CC	be used in gene therapy. HSPPs can be used to treat or prevent disorders
CC	associated with decreased activity or function of HSP. Antagonists of
CC	HSPP are used to treat or prevent disorders associated with increased
CC	activity or function of HSPP. Such diseases include cell proliferation
CC	(including cancer), inflammation, cardiovascular, neurological,
CC	reproductive or developmental disorders, (e.g. arteriosclerosis,
CC	cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,
CC	asthma, Crohn's disease, microbial or other infections, congestive or
CC	ischemic heart disease, Alzheimer's, Parkinson's or Huntington's
CC	diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSPP
CC	nucleic acids can be used for the recombinant production of HSPP, for
CC	detecting HSPP in standard hybridisation and amplification assays (for
CC	diagnosis and monitoring), in gene therapy, as antisense,
CC	triplex-forming or ribozyme therapeutics, for detecting related sequences
CC	or genetic variations, and for chromosomal mapping. HSPP are also used to
CC	raise specific antibodies (Ab) and to screen for agonists and
CC	antagonists (potential therapeutic agents). Ab are used to diagnose, or
CC	monitor, HSPP-related diseases (in usual immunoassays), as therapeutic
CC	antagonists, in competitive drug screens, and for purification of HSPP
CC	from natural sources.
XX	
SQ	Sequence 748 BP; 159 A; 227 C; 193 G; 169 T; 0 other;
	Alignment Scores:
	Pred. No.: Length: 748
	Score: 661.00 Matches: 116
	Percent Similarity: 100.00% Conservative: 0
	Best Local Similarity: 100.00% Mismatches: 0
	Query Match: 98.95% Indels: 0
	DB: 21 Gaps: 0
US-09-092-297-17 (1-117) x AAZ98158 (1-748)	
OY	2 leuGlnProProkroArgAlaMetAlaProArgGlyCysIleValAlaValAPheAlaIlePhe 21
Dd	
	1 CTGCACACCACCGAGCGCATGGCTCCCCGAGGCTGCATCGTAGTGTCTTTGCCATTTC 60
OY	22 CysIleSerArgLeuLeuCySserHISglYAlaProValAlaProMetTrnProTyrlieu 41
Dd	
	61 TGGAATCTCACGGCTCTCTGCTCAACAAGGAGCCCAATGGGCCCCAATGACTCTTACTCG 120
OY	42 MetLeuCysGlnProHisIstLysArgCysGlyAspLysTherYrAspProLeuGlnHisCys 61
Dd	
	121 ATGCTTGTCACGCACACAGAGATGTGGGACAAAGTTTCAACGACCCCCTGCACACACTGT 180

Oy	62	CysGlyrAspAspAlaValAlaProLeuAlaArgThrGlnThCysGlyAsnGlyThrPhe	81
Db	181	TGCTATGATGATGGCTGCTGTCCTTGGCCAGACCCAGACGCTGTGGAAATGCACCTTC	240
Oy	82	ArgValCysPheGluGlnCysCysProTyrPThrPheMetValIysLeuIleAsnIasn	101
Db	241	AGAATCTGCTTTGAGCAGTGCTGCTGCCCTGACCTTCATGTTGAAGCTGATTAACCGAAGC	300
Oy	102	CysAspSerAlaArgThrSerAspArgLeuCysArgSerValSer	117
Db	301	TGGCAGTCACGCCCGACCTCGAGATGACAGACGCTTGTGCGCAGTGTACGC	348
RESULT 11			
AXX01684	AAAX01684 standard; DNA; 236 BP.		
XX	AAAX01684:		
XX	21-APR-1999	(first entry)	
DT	Urinary tract tissue library BL172 gene specific cDNA clone 2624118.		
DE	BL172; urinary tract; marker; cancer; recombinant; human; ss.		
XX	Homo sapiens.		
OS	Homo sapiens.		
PN	W09855656-A1.		
XX	10-DEC-1998.		
PD	05-JUN-1998;	98WO-US11693.	
PF	05-JUN-1997;	97US-0869579.	
PR	(ABBO) ABBOTT LAB.		
XX	Billing-medel PA, Cohen M, Colpitts TL, Friedman PN;		
PI	Gordon J, Grandosen, Hodges SC, Klass MR, Kratochvil JD;		
PI	Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;		
XX	WPI: 1999-045802/04.		
DR	New purified polynucleotide BL172 derivatives and encoded		
PT	polypeptides - useful in the identification of markers, which are		
PT	indicative of urinary tract diseases or conditions		
XX	Claim 1; Page 89; 114pp: English.		
PS	Sequences AAX01683-85 represent overlapping clones specific for urinary		
XX	tract tissue library BL172. The BL172 polynucleotides are used to detect		
CC	target BL172 polynucleotides and BL172 mRNA and the BL172 polypeptides		
CC	are used to detect BL172 antigens/antibodies in a test sample. The		
CC	identification of certain markers in these methods are indicative of the		
CC	presence of urinary tract disease, especially cancer. Additionally, the		
CC	polypeptides are used to detect antibodies that bind specifically to a		
CC	BL172 epitope derived from a BL172 polypeptide. Host cells containing a		
CC	recombinant expression vector comprising the BL172 polynucleotide		
CC	sequences are used to produce the polypeptides containing BL172		
CC	epitopes. The methods aid in diagnosis, detection, staging, monitoring,		
CC	prognostication, in vivo imaging, prevention and treatment of diseases		
CC	or conditions associated with BL172, especially urinary tract cancer.		
XX	Sequence 236 BP; 49 A; 69 C; 62 G; 56 T; 0 other;		
SO			
Alignment Scores:			
Pred. No.:	1.57e-32	Length:	236
Score:	372.00	Matches:	64
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	55.69%	Indels:	0
DB:	20	Gaps:	0

QY 54 PheTyAspProLeuGlnHisCysCysTyrrAspAlaValAlaProLeuAlaArgThr 73
 Db 1 TCTTCAGACCCCTCGACAGACTGTGCTATGATGATGCCGTGCTGCTGGCCAGAC 60
 QY 74 GlnThrCysGlyAsnCysThrPheArgValCysPheGluGlnCysCysProThrPhe 93
 Db 61 CAGAGCTGTGGAACTGCACCTTCAGAGTCTGCTTGAAGCAGTGTGCTGCCCTTC 120
 QY 94 MetValLysLeuLeuAsnGlnAsnCysAspSerAlaArgThrSerAspArgLeuCys 113
 Db 121 ATGGGAGAGCTGATTAACACCACTGCACCTACCGCCGAGCTCGATGACAGCTTGT 180
 QY 114 ArgSerValSer 117
 Db 181 CGCAGTGTCAAGC 192

RESULT 12
 AAX01683
 ID AAX01683 standard; DNA; 196 BP.
 AC AAX01683;
 DT 21-APR-1999 (first entry)
 DE Urinary tract tissue library BL172 gene specific cDNA clone 1554838.
 KW BL172; urinary tract; marker; cancer; recombinant; human; ss.
 OS Homo sapiens.
 PN W0985656-A1.
 PD 10-DEC-1998.
 PF 05-JUN-1998; 98WO-US11693.
 PR 05-JUN-1997; 97US-0869579.
 PA (ABBO) ABBOTT LAB.
 PI Billing-medel PA, Cohen M, Colpitts TL, Friedman PN;
 PI Gordon J, Grandosen, Hodges SC, Klass MR, Kratochvil JD;
 PI Roberts-Kapp L, Russell JC, Stroupe SD, Yu H;
 DR WPI; 1999-045802/04.
 XX New purified polynucleotide BL172 derivatives and encoded
 PT polypeptides - useful in the identification of markers, which are
 PS indicative of urinary tract diseases or conditions.
 Claim 1; Page 89; 114pp; English.

Sequences AAX01683-85 represent overlapping clones specific for urinary tract tissue library BL172. The BL172 polynucleotides are used to detect target BL172 polynucleotides and BL172 mRNA and the BL172 polypeptides are used to detect BL172 antigens/antibodies in a test sample. The identification of certain markers in these methods are indicative of the presence of urinary tract disease, especially cancer. Additionally, the polypeptides are used to detect antibodies that bind specifically to a BL172 epitope derived from a BL172 polypeptide. Host cells containing a recombinant expression vector comprising the BL172 polynucleotide sequences are used to produce the polypeptides containing BL172 epitopes. The methods aid in diagnosis, detection, staging, monitoring, or conditions associated with BL172, especially urinary tract cancer.

Alignment Scores:
 Pred. No.: 2.05e-32 Length: 196
 Score: 370.00 Matches: 64
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 55.39% Indels: 0
 DB: 20 Gaps: 0

US-09-092-297-17 (1-117) x AAX01683 (1-196)

QY 1 ProluGlnProProArgAlaMetAlaProArgGlyCysIleValAlaValPheAlaIle 20
 Db 3 CCACGCAACACCCAGAGCCATGCTCCCGAGGCTCCTCAGTGTGCTTTCAT 62
 QY 21 PheCysIleSerArgLeuLeuCysSerHisGlyAlaProValAlaProMetThrProTyr 40
 Db 63 TTTGCTGCTTCAGGCTCTCTCTCTCACACGAGCCGCCAGTGGCCCATCTCTTAC 122
 QY 41 LeuMetLeuGlnProHisLysArgCysGlyAspLysPheTyrrAspProLeuGlnHis 60
 Db 123 CTGATGCTGTGCCAGCCACACAGAGATGTGGGCAAGATTCTAGACCCCTTCACAC 182
 QY 61 CysCysTyrrAsp 64
 Db 183 TGTGCTATCAT 194

RESULT 13
 AAS92056
 ID AAS92056 standard; CDNA; 1338 BP.
 AC AAS92056;
 DT 13-FEB-2002 (first entry)
 DE DNA encoding novel human diagnostic protein #27860.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 PN W0200175067-A2.
 PD 11-OCT-2001.
 PF 30-MAR-2001; 2001WO-US08631.
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 DR WPI; 2001-639362/73.
 DR P-PSDB; ABG27869.

New isolated polynucleotide and encoded polypeptides, useful in diagnosis, forensics, gene mapping, identification of mutations biodiversity -

Claim 1; SEQ ID No 27860; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in

Score: NO.:	4,86e-17	Length:	662
Percent Similarity:	237.00	Matches:	49
Best Local Similarity:	55.008	Conservative:	17
Query Match:	40.838	Mismatches:	34
DB:	35.488	Indels:	20
	21	Gaps:	3

05-09-092-297-17 (1-117) x AAC58632 (1-662)

Oy		7	AlaMetAlaProArgGlyCysIleValAlaValPheAlaIlePheCysIleSerArgLeu	26
Db		23	ACCAATAGAGCCACGATGCATCTTGCTTGTCTGGATAAACAGTCCTCTCCTC	82
Oy		27	LeuCysSerHisGLYAlaProValAlaProMetThrProTyrlLeuMetLeuCysGlnPro	46
Db		83	CAGTTCTTAAGAAGAACACTACAGACGCCCTCTTGTGGTCACAGACACTGTGGCTGGCACCG	142
Oy		47	HistylsArGcysglAspLysPheTyraSPProLeuGlnHiscysCytryrasPasPaLa	66
Db		143	ACACCACGGTGTGGGAACAAGANCTACAAACCCCTTCACAGACGACGCTGTATGATGATGCC	202
Oy		67	ValValProleuAlaAlythrGlnInrthrcysgLyAsn---CysthrPharGvalCysPhe	85
Db		203	ATCTTATCTTAAAGAGACCCGCCCTGTGGCTTCACACTGCACCTTCTGGCCCTGCTTT	262
Oy		86	GluGlnAcscyscysProTrpThr-----PheMetValLysLeu-----	97
Db		263	GACCTCTGCTGTGCCAGACTTTTGGCCCCCAGACAGAAATTCTTGTGTAAGTTGAGGCTT	322
Oy		98	-----IleanglnAsnCysAspSeralAarg	106
Db		323	CTGGTAGTAGCTCATGTCAGTGCATTATCTCCATCTCCGAGCTTGACAGAAACAG	382
RESULT 15				
ID	AA51261			
XX	AA51261 standard; cDNA; 662 bp.			
AC	AA51261;			
XX				
DT	26-SEP-2000 (first entry)			
XX				
DE	Human DNA encoding PRO982, a novel secreted protein.			
KW	PRO982; secreted protein; transmembrane protein; recombinant production			
XX	gene therapy; ss.			
OS	Homo sapiens.			
FH				
FT	Key	Location/Qualifiers		
FT	CDS	26..403		
FT		/*tag= a		
FT		/product= PRO982		
FT	sig_peptide	/note= "a secreted protein"		
XX		26..88		
PD	WO200036102-A2.	/*tag= b		
XX				
PF	01-DEC-1999;	99WO-US28634.		
PR	16-DEC-1998;	98US-0112851.		
PR	16-DEC-1998;	98US-0113145.		
PR	22-DEC-1998;	98US-0113511.		
PR	12-JAN-1999;	99US-0115558.		
PR	12-JAN-1999;	99US-0115565.		
PR	12-JAN-1999;	99US-0115733.		
PR	09-FEB-1999;	99US-0119341.		
PR	10-FEB-1999;	99US-0119537.		
PR	12-FEB-1999;	99US-0119565.		

PR 02-JUN-1999; 99MO-US12252
XX

(GETH) GENENTECH INC.
PA
XX

P1	Gurney AL,	Pan J,	Wen X
P1	Wood WI;	Roy MA,	Wen X
PI		Stewart TA,	Wen X
XX		Tumas D,	Wen X
		Watanabe CK;	Wen X
		Gao W,	Wen X
		Goddard A,	Wen X
		Fong S,	Wen X
		Watanabe N,	Wen X

DR P-PSDB; AAY96731.

PT isolated nucleic acid molecule encodes a PRO polypeptide which is a transmembrane polypeptide

Claim 1; Fig. 5; 154pp; English.

CC This cDNA encodes PRO982, a novel secreted protein. The invention
CC concerns novel secreted and transmembrane proteins, designated PRO
CC polypeptides. The cDNA and gene sequences are useful in the recombinant
CC production of PRO polypeptides, as a hybridization probe to screen
CC libraries to isolate cDNAs with sequence identity to PRO polypeptides or
CC to map the gene encoding the PRO polypeptides and analyzing genetic
CC disorders. The cDNA/gene can also be used to produce transgenic animals
CC useful for the development and screening of therapeutically useful
CC reagents. They can also be used in gene therapy, e.g. to replace a
XX defective gene.

Sequence 662 BP; 152 A; 187 C; 167 G; 156 T; 0 other;
Alignment Score: 100

Score:	4.86e-17	Length:	662
Percent Similarity:	237.00	Matches:	49
Best Local Similarity:	55.0%	Conservative:	17
Query Match:	40.83%	Mismatches:	34
DB:	35.488	Indels:	20
	21	Gaps:	2

03-092-291-17 (1-117) x AA51261 (1-662)

[illegible]

Search completed: November 3, 2002, 03:06:00
Job time : 213 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002. CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 23:14:12 ; Search time 209 Seconds

(without alignments)
6267.970 Million cell updates/sec

Title: US-09-092-297-5

Perfect score: 763
Sequence: 1 CTCCTGACACACACACACAG.....ATTAATTATGACTTTATA 763

Scoring table: OLIGO-NUC

Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802.*

```
1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	763	100.0	763	20	AAAX01687
2	712	93.3	763	20	AAAX01686
3	710	93.1	762	20	AAZ24391
4	661	86.6	771	21	AAAZ7668
5	661	86.6	771	21	AAAZ7062
6	661	86.6	771	22	AAZ21491
7	661	86.6	771	22	AAZ54298
8	661	86.6	914	21	AAAO8111
9	661	86.6	914	21	AAAO8111

10	569	74.6	748	21	AAZ98158
C 11	366	48.0	417	20	AAAX01685
12	236	30.9	236	20	AAAX01684
13	196	25.7	196	20	AAAX01683
14	23	3.0	1338	23	AAZ92056
15	22	2.9	22	20	AAAX01690
C 16	22	2.9	22	20	AAAX01691
C 17	21	2.8	48	22	AAZ5970
C 18	21	2.8	9229	9	AAZ80437
C 19	20	2.6	47	22	AAZ5971
C 20	19	2.5	338	21	AAZ5372
21	19	2.5	423	18	AAZ80723
22	19	2.5	2421	19	AAZ9808
23	19	2.5	2421	21	AAZ00747
24	19	2.5	2487	19	AAZ00748
25	18	2.4	242	21	AAZ05699
C 26	18	2.4	252	20	AAZ98548
C 27	18	2.4	252	20	AAZ98548
28	18	2.4	307	22	AAZ55566
29	18	2.4	422	18	AAZ80646
C 30	18	2.4	477	22	AAZ54027
C 31	18	2.4	477	22	AAZ02300
C 32	18	2.4	477	22	AAZ27744
33	18	2.4	491	22	AAZ65324
34	18	2.4	508	22	AAZ98095
C 35	18	2.4	714	22	AAZ16681
C 36	18	2.4	714	22	AAZ16682
C 37	18	2.4	714	22	AAZ16683
38	18	2.4	714	22	AAZ79185
39	18	2.4	714	22	AAZ79186
40	18	2.4	714	22	AAZ79189
41	18	2.4	2844	20	AAZ87842
42	18	2.4	3156	20	AAZ73917
43	18	2.4	3156	20	AAZ71603
44	18	2.4	3156	20	AAZ69285
45	18	2.4	4596	18	AAZ03526

ALIGNMENTS

RESULT 1	AAAX01687	standard; DNA; 763 BP.
ID	AAAX01687	
AC	AAAX01687	
XX	21-APR-1999	(first entry)
DE	Urinary tract tissue library BL172 gene consensus sequence.	
XX	BL172; urinary tract; marker; cancer; recombinant; human; ss.	
KW	Homo sapiens.	
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	3..356
FT		/tag= a
FT		/product= "BL172 polypeptide"
FT		/note= "the start codon is not indicated"
XX		
PN	W09855656-A1.	
XX		
PD	10-DEC-1998.	
XX		
PF	05-JUN-1998;	98WO-US11693.
XX		
PR	05-JUN-1997;	97US-0869579.
XX		
PA	(ABBO) ABBOT LAB.	
XX		
PI	Billings-medell PA, Cohen M, Colpitts TL, Friedman PN, Gordon J, Granadosen, Hodges SC, Klass MR, Kratochvil JD;	

PI Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;
 XX MPI, 1999-045802/04.
 DR P-PSDB; AAW92252.
 XX
 PT New purified polynucleotide BL172 derivatives and encoded
 PT polypeptides - useful in the identification of markers, which are
 PT indicative of urinary tract diseases or conditions
 XX
 PS Claim 1: Page 90; 114pp; English.
 XX
 CC This represents a consensus sequence of a clone specific for urinary
 CC tract tissue library BL172. The BL172 polynucleotides (AA01683-87) are
 CC used to detect target BL172 polynucleotides and BL172 mRNA and the BL172
 CC polypeptides are used to detect BL172 antigens/antibodies in a test
 CC sample. The identification of certain markers in these methods are
 CC indicative of the presence of urinary tract disease, especially cancer.
 CC Additionally, the polypeptides are used to detect antibodies that bind
 CC specifically to a BL172 epitope derived from a BL172 polypeptide. Host
 CC cells containing a recombinant expression vector comprising the BL172
 CC polynucleotide sequences are used to produce the polypeptides containing
 CC BL172 epitopes. The methods aid in diagnosis, detection, staging,
 CC monitoring, prognosis, prevention, in vivo imaging, prevention and treatment
 CC of diseases or conditions associated with BL172, especially urinary
 CC tract cancer.
 XX
 XX Sequence 763 BP; 162 A; 231 C; 194 G; 175 T; 1 other;
 SQ
 Query Match 100.0%; Score 763; DB 20; Length 763;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 763; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTACTGCAACCCAGACGAGCCAGTGGCTCCCGAGGCTGATCGTGTCTTTGCCA 60
 DB 1 CTCCTACTGCAACCCAGACGAGCCAGTGGCTCCCGAGGCTGATCGTGTCTTTGCCA 60
 QY 61 TTTTGTGATCTCGAGGCTCTGTCTGACAGGAGCCCGGCTGATCGTGTCTTTGCCA 120
 DB 61 TTTTGTGATCTCGAGGCTCTGTCTGACAGGAGCCCGGCTGATCGTGTCTTTGCCA 120
 QY 121 ACTGTGCTGTGCGACGACACAGAGATGTGGGACAACTTCTACGACCCCTGACAGC 180
 DB 121 ACTGTGCTGTGCGACGACACAGAGATGTGGGACAACTTCTACGACCCCTGACAGC 180
 QY 181 ACTGTGCTATGATGATGCGCTGCTGCTTGGCCAGACCCAGAGCTGTGAAATGCA 240
 DB 181 ACTGTGCTATGATGATGCGCTGCTGCTTGGCCAGAGCGTGTGAAATGCA 240
 QY 241 CCTTACAGTCTGCTTGGAGAGTGTGCTGCGCTTGTGAGCTTGTGAAATGCAATACC 300
 DB 241 CCTTACAGTCTGCTTGGAGAGTGTGCTGCGCTTGTGAGCTTGTGAAATGCAATACC 300
 QY 301 AGAAGTGGAGCTCAACCCGAGCTGATGACAGGCTTGTGCGAGTGTACATATGGA 360
 DB 301 AGAAGTGGAGCTCAACCCGAGCTGATGACAGGCTTGTGCGAGTGTACATATGGA 360
 QY 361 ACATAGGGGAGAGATGATGCTGATGCTCTCTCTGAGGCTGAGGAGAAAGAGCT 420
 DB 361 ACATAGGGGAGAGATGATGCTGATGCTCTCTCTGAGGCTGAGGAGAAAGAGCT 420
 QY 421 GGTGTACCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
 DB 421 GGTGTACCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
 QY 481 AACTGCCCCATCTTCTGTGACGCTGTGAGGCCACCTGCGCTGCTGAGAGAGC 540
 DB 481 AACTGCCCCATCTTCTGTGACGCTGTGAGGCCACCTGCGCTGCTGAGAGAGC 540
 QY 541 CCACAGGTCCTCTCTAAGTCTGACAGATGATGATGATGATGATGATGATGATGATG 600
 DB 541 CCACAGGTCCTCTCTAAGTCTGACAGATGATGATGATGATGATGATGATGATGATG 600
 QY 601 GGAATCTGAACCTCTCTGATGACCCSTATGGCAACATCAACCCGGACACCCCAAGC 660

DB 601 GGAATCTGAACCTCTCTGATGACCCSTATGGCAACATCAACCCGGACACCCCAAGC 660
 QY 661 TGGCTGGGAGACCCCTTCCAGGATTTTCCATCATCATCATCATCATCATCATCATCAT 720
 DB 661 TGGCTGGGAGACCCCTTCCAGGATTTTCCATCATCATCATCATCATCATCATCATCAT 720
 QY 721 CAGAGCAAAAGCAGAGATCATTAATAATTATGACTTTATA 763
 DB 721 CAGAGCAAAAGCAGAGATCATTAATAATTATGACTTTATA 763
 RESULT 2
 AA01686
 ID AA01686 standard; DNA; 763 BP.
 AC AA01686;
 XX
 DT 21-APR-1999 (first entry)
 XX
 DE Urinary tract tissue library BL172 gene full length clone 1554838H.
 XX
 KW BL172; urinary tract; marker; cancer; recombinant; human; ss.
 XX
 OS Homo sapiens.
 XX
 PN W09855656-A1.
 XX
 PD 10-DEC-1998.
 XX
 PF 05-JUN-1998; 98WO-US11693.
 XX
 PR 05-JUN-1997; 97US-0869579.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 PI Billing-medell PA, Cohen M, Colplitts TL, Friedman PN;
 PI Gordon J, Granadosen, Hodges SC, Klaas MR, Kratochvil JD;
 PI Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;
 DR MPI, 1999-045802/04.
 XX
 PT New purified polynucleotide BL172 derivatives and encoded
 PT polypeptides - useful in the identification of markers, which are
 PT indicative of urinary tract diseases or conditions
 XX
 PS Claim 1: Page 89-90; 114pp; English.
 XX
 CC This represents a full-length sequence of a clone specific for urinary
 CC tract tissue library BL172. The BL172 polynucleotides (AA01683-87) are
 CC used to detect target BL172 polynucleotides and BL172 mRNA and the BL172
 CC polypeptides are used to detect BL172 antigens/antibodies in a test
 CC sample. The identification of certain markers in these methods are
 CC indicative of the presence of urinary tract disease, especially cancer.
 CC Additionally, the polypeptides are used to detect antibodies that bind
 CC specifically to a BL172 epitope derived from a BL172 polypeptide. Host
 CC cells containing a recombinant expression vector comprising the BL172
 CC polynucleotide sequences are used to produce the polypeptides containing
 CC BL172 epitopes. The methods aid in diagnosis, detection, staging,
 CC monitoring, prognosis, prevention, in vivo imaging, prevention and treatment
 CC of diseases or conditions associated with BL172, especially urinary
 CC tract cancer.
 XX
 XX Sequence 763 BP; 162 A; 231 C; 195 G; 175 T; 0 other;
 SQ
 Query Match 93.3%; Score 712; DB 20; Length 763;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 762; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCCTACTGCAACCCAGACGAGCCAGTGGCTCCCGAGGCTGATCGTGTCTTTGCCA 60
 DB 1 CTCCTACTGCAACCCAGACGAGCCAGTGGCTCCCGAGGCTGATCGTGTCTTTGCCA 60

QY 61 TTTTCGATCTCCAGGCTCCTCTGTCTACACGAGAGCCCGAGTGGCCCGCCATGACTCCT 120
 DB 61 TTTTCGATCTCCAGGCTCCTCTGTCTACACGAGAGCCCGAGTGGCCCGCCATGACTCCT 120
 QY 121 ACCTGATGCTGTGCGCCACACACAGAGATGTGGGACAGATTCTACGACCCCTGAGC 180
 DB 121 ACCTGATGCTGTGCGCCACACACAGAGATGTGGGACAGATTCTACGACCCCTGAGC 180
 QY 181 ACCTGATGCTGTGCGCCACACACAGAGATGTGGGACAGATTCTACGACCCCTGAGC 240
 DB 181 ACCTGATGCTGTGCGCCACACACAGAGATGTGGGACAGATTCTACGACCCCTGAGC 240
 QY 241 CCTTCAGAGTCTGCTGTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 DB 241 CCTTCAGAGTCTGCTGTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 QY 301 AGAATGCGACTAGCCCGGACCTGGAGTACAGAGCTTTGTGCGAGTGTGAGTAAATGA 360
 DB 301 AGAATGCGACTAGCCCGGACCTGGAGTACAGAGCTTTGTGCGAGTGTGAGTAAATGA 360
 QY 361 ACATCAGGGGAGAGATGACTCCCTGAGATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 DB 361 ACATCAGGGGAGAGATGACTCCCTGAGATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 421 GGTGTACCTGAGATCTGGGATGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 DB 421 GGTGTACCTGAGATCTGGGATGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 481 AACTGCCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 DB 481 AACTGCCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 QY 541 CCACAGAGTCCCTCTGAGATTTGTGACAGATGAGATGCTGCTGCTGCTGCTGCTGCTGCT 600
 DB 541 CCACAGAGTCCCTCTGAGATTTGTGACAGATGAGATGCTGCTGCTGCTGCTGCTGCTGCT 600
 QY 601 GGACTGTGACCTCTGATGACCCCTGATGACCCCTGATGACCCCTGATGACCCCTGATGAC 660
 DB 601 GGACTGTGACCTCTGATGACCCCTGATGACCCCTGATGACCCCTGATGACCCCTGATGAC 660
 QY 661 TGGCTGGGAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 DB 661 TGGCTGGGAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 721 CAGGAGCAAGACAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 763
 DB 721 CAGGAGCAAGACAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 763
 RESULT 3
 ID AA24391 standard; cDNA; 762 BP.
 AC AA24391;
 DT 14-FEB-2000 (first entry)
 XX Human bladder tumour cDNA library derived EST 3.
 DE Human bladder tumour cDNA library derived EST 3.
 XX Expressed sequence tag; human; bladder; tumour; cancer; cytostatic;
 KW treatment; gene therapy; EST; ss.
 XX Homo sapiens.
 OS
 XX DE19818619-A1.
 XX 28-OCT-1999.
 XX 21-APR-1998; 98DE-1018619.
 XX 21-APR-1998; 98DE-1018619.
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX

XX
 PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
 XX WPI; 1999-612028/53.
 DR
 XX
 PT New nucleic acid sequences expressed in bladder tumor tissue, and
 PT derived polypeptides, for treatment of bladder tumor and identification
 of therapeutic agents
 XX
 PS Claim 3; Page 62; 132pp; German.
 CC
 CC This invention describes novel polypeptide fragments (I) and the
 CC polynucleotides (II) that encode them that are highly expressed in a
 CC human bladder tumour and which have cytostatic activity. (II) are used
 CC for recombinant expression of (I) and to isolate complete genes. (I) are
 CC used to identify agents suitable for treatment of bladder cancer. (I)
 CC directly treat this form of cancer (including expression from gene
 CC therapy vectors) or are used in a preparation for cancer treatment. (I)
 CC is also used for the generation of specific antibodies. (II) are
 CC identified by assembling ESTs (expressed sequence tags) from a
 CC particular tissue type before comparison of expression patterns. This
 CC allows a significantly longer fragment of the gene to be revealed, and
 CC therefore reduces the number of failures associated with the fact that
 CC ESTs from different libraries may represent different parts of the same
 CC unknown gene, distorting the estimated frequency of occurrence in a
 CC particular tissue. AA24391-243109 represent expressed sequence tag (EST)
 CC fragments isolated from a human bladder tumour cDNA library which encode
 CC the proteins represented in AA24391-243109.
 XX
 SQ Sequence 762 BP; 162 A; 232 C; 194 G; 174 T; 0 other;
 Query Match 93.1%; Score 710; DB 20; Length 762;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 760; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CTCGATGCAACACCCAGAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
 DB 1 CTCGATGCAACACCCAGAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
 QY 61 TTTTCGATCTCCAGGCTCCTCTGTCTACACGAGAGCCCGAGTGGCCCGCCATGACTCCT 120
 DB 61 TTTTCGATCTCCAGGCTCCTCTGTCTACACGAGAGCCCGAGTGGCCCGCCATGACTCCT 120
 QY 121 ACCTGATGCTGTGCGCCACACACAGAGATGTGGGACAGATTCTACGACCCCTGAGC 180
 DB 121 ACCTGATGCTGTGCGCCACACACAGAGATGTGGGACAGATTCTACGACCCCTGAGC 180
 QY 181 ACTGTGCTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 DB 181 ACTGTGCTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 QY 241 CCTTCAGAGTCTGCTGTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 DB 241 CCTTCAGAGTCTGCTGTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 QY 301 AGAATGCGACTAGCCCGGACCTGGAGTACAGAGCTTTGTGCGAGTGTGAGTAAATGA 360
 DB 301 AGAATGCGACTAGCCCGGACCTGGAGTACAGAGCTTTGTGCGAGTGTGAGTAAATGA 360
 QY 361 ACATCAGGGGAGAGATGACTCCCTGAGATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 DB 361 ACATCAGGGGAGAGATGACTCCCTGAGATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 421 GGTGTACCTGAGATCTGGGATGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 DB 421 GGTGTACCTGAGATCTGGGATGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 481 AACTGCCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 DB 481 AACTGCCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 QY 541 CCACAGAGTCCCTCTGAGATTTGTGACAGATGAGATGCTGCTGCTGCTGCTGCTGCTGCT 600
 DB 541 CCACAGAGTCCCTCTGAGATTTGTGACAGATGAGATGCTGCTGCTGCTGCTGCTGCTGCT 600

Db 541 CCACAGTCCCTTCTAGAAATCTGACAGCATGAGTGGTGTGATGGGGGCCAG 600
 QY 601 GGACTGTGAACCTCTCTGATGACCCCTATGGCCAAATCAACCCGGCCACCCCAAGGC 660
 Db 601 GGACTGTGAACCTCTCTGATGACCCCTATGGCCAAATCAACCCGGCCACCCCAAGGC 660
 QY 661 TGACCTGGGGAACCTTCAACCTCTCTGATGATTTTCATCATCTCAAGTTCCTTCTATC 720
 Db 661 TGACCTGGGGAACCTTCAACCTCTCTGATGATTTTCATCATCTCAAGTTCCTTCTATC 720
 QY 721 CAGGAGCAAGACAGCATCATTAATAATTATGACTTTA 761
 Db 721 CAGGAGCAAGACAGCATCATTAATAATTATGACTTTA 761
 RESULT 4
 AAA77668
 ID AAA77668 standard; cDNA; 771 BP.
 AC AAA77668;
 XX 07-NOV-2000 (first entry)
 DE Human FRO1274 cDNA sequence SEQ ID NO:196.
 XX
 KW Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;
 KW diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;
 KW angiogenic; proliferative; cardiac; cardiovascular; antiatherosclerotic;
 KW cytosolic; gene therapy; vaccine; ss.
 XX
 OS Homo sapiens.
 PN WO200032221-A2.
 PD 08-JUN-2000.
 XX
 PF 30-NOV-1999; 99MO-US28313.
 PR 01-DEC-1998; 98MO-US25108.
 PR 16-DEC-1998; 98US-0112850.
 PR 12-JAN-1999; 99US-0115554.
 PR 08-MAR-1999; 99MO-US05028.
 PR 12-MAR-1999; 99US-0123957.
 PR 28-APR-1999; 99US-0131445.
 PR 14-MAY-1999; 99US-0134287.
 PR 02-JUN-1999; 99MO-US12522.
 PR 23-JUN-1999; 99US-0141037.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.
 PR 01-SEP-1999; 99MO-US20111.
 PR 08-SEP-1999; 99MO-US20594.
 PR 13-SEP-1999; 99MO-US20944.
 PR 15-SEP-1999; 99MO-US21090.
 PR 15-SEP-1999; 99MO-US21547.
 PR 05-OCT-1999; 99MO-US23089.
 PR 29-OCT-1999; 99US-0162506.
 XX
 PA (GETH) GENENTECH INC.
 PI Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hslian KJ, Goddard A;
 PI Godowski PJ, Gurney AL, Klein RD, Kuo SS, Paoni NF, Smith V;
 PI Watanabe CK, Williams PM, Wood WI;
 DR WPI: 2000-412154/35.
 DR P-PSDB; AAB24425.
 XX
 PT Nucleic acids encoding PRO polypeptides useful for preventing,
 PT diagnosing and treating disorders in mammals -
 PT angiogenic disorders in mammals -
 XX Claim 61; Fig 75; 315pp; English.
 XX The present invention describes nucleic acids encoding PRO polypeptides

CC useful for preventing, diagnosing and treating disorders in mammals by
 CC cardiovascular, endothelial or angiogenic disorder in mammals by
 CC modulating cell proliferation, angiogenesis and cardiovascularisation,
 CC and for identifying agonists and antagonists of these processes. The
 CC nucleic acids and the proteins they encode may be used in the
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate PRO expression such as cardiovascular, endothelial or
 CC angiogenic disorders in mammals (e.g. atherosclerosis, cancers and
 CC cardiac hypertrophy). For example, the nucleic acids (NCs) and vectors
 CC associating them and the PRO polypeptide may be used to treat disorders
 CC associated with decreased PRO expression. AAA77510 to AAA77721 and
 CC AAB24388 to AAB24435 represent nucleotide and protein sequences used in
 CC the exemplification of the present invention.
 XX

Sequence 771 BP; 169 A; 231 C; 195 G; 176 T; 0 other;

Query Match 86.6%; Score 661; DB 21; Length 771;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 761; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTCACACTGCACACCCAGAGCATGCTCCCGAGGCTGATCGTAGCTGCTTGGCA 60
 Db 1 CTCACACTGCACACCCAGAGCATGCTCCCGAGGCTGATCGTAGCTGCTTGGCA 60
 QY 61 TTTTCTGCATCTCCAGGCTCTCTGCTACACAGGAGCCCAAGTGGCCCATGACTCCTT 120
 Db 61 TTTTCTGCATCTCCAGGCTCTCTGCTACACAGGAGCCCAAGTGGCCCATGACTCCTT 120
 QY 121 ACCTGATGCTGGCCAGCCACAGAGAGATGAGGAGCAAGTTCTACAGCCCTCTACG 180
 Db 121 ACCTGATGCTGGCCAGCCACAGAGAGATGAGGAGCAAGTTCTACAGCCCTCTACG 180
 QY 181 ACTGTTGCTATGATGATGCGCTGCTGCTGCTGCGCCAGAGCCAGAGCTGTGAAATGCA 240
 Db 181 ACTGTTGCTATGATGATGCGCTGCTGCTGCTGCGCCAGAGCCAGAGCTGTGAAATGCA 240
 QY 241 CCTTCAGAGTCTGCTTTGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 Db 241 CCTTCAGAGTCTGCTTTGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 QY 301 AGAAGTGGAGTACAGCCCGGAGCTGGAGAGACAGGTTTGTGCTGCTGCTGCTGCTGCT 360
 Db 301 AGAAGTGGAGTACAGCCCGGAGCTGGAGAGACAGGTTTGTGCTGCTGCTGCTGCTGCT 360
 QY 361 ACATCAGGGAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Db 361 ACATCAGGGAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 421 GGTGTACCTGAGATCTGGAGTCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 Db 421 GGTGTACCTGAGATCTGGAGTCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 481 AACTGCCCATCTATCTGTGACCTGTGTGAGGCCACCCCTCAGCTGCTGAGAGAGC 540
 Db 481 AACTGCCCATCTATCTGTGACCTGTGTGAGGCCACCCCTCAGCTGCTGAGAGAGC 540
 QY 541 CCACAGTCCCTTCTAGAAATCTGACAGCATGAGTGGTGTGATGGGGGCCAG 600
 Db 541 CCACAGTCCCTTCTAGAAATCTGACAGCATGAGTGGTGTGATGGGGGCCAG 600
 QY 601 GGACTGTGAACCTCTCTGATGACCCCTATGGCCAAATCAACCCGGCCACCCCAAGGC 660
 Db 601 GGACTGTGAACCTCTCTGATGACCCCTATGGCCAAATCAACCCGGCCACCCCAAGGC 660
 QY 661 TGACCTGGGGAACCTTCAACCTCTCTGATGATTTTCATCATCTCAAGTTCCTTCTATC 720
 Db 661 TGACCTGGGGAACCTTCAACCTCTCTGATGATTTTCATCATCTCAAGTTCCTTCTATC 720
 QY 721 CAGGAGCAAGACAGCATCATTAATAATTATGACTTTA 763
 Db 721 CAGGAGCAAGACAGCATCATTAATAATTATGACTTTA 763

RESULT 5
AAA37062
ID AAA37062 standard; cDNA: 771 BP.
XX
AC AAA37062;
XX
DT 08-AUG-2000 (first entry)
XX
DE Human PRO1274 (UNQ644) cDNA sequence SEQ ID NO:137.
XX
KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
transmembrane; secretion; immunoadhesion; pharmaceutical; screening;
ss.
XX
OS Homo sapiens.
XX
PN W0200012708-A2.
XX
PD 09-MAR-2000.
XX
PF 01-SEP-1999; 99WO-US20111.
XX
XX 01-SEP-1998; 98US-0098716.
PR 01-SEP-1998; 98US-0098749.
PR 01-SEP-1998; 98US-0098750.
PR 02-SEP-1998; 98US-0098803.
PR 02-SEP-1998; 98US-0098821.
PR 02-SEP-1998; 98US-0098843.
PR 09-SEP-1998; 98US-0099536.
PR 09-SEP-1998; 98US-0099596.
PR 09-SEP-1998; 98US-0099598.
PR 09-SEP-1998; 98US-0099602.
PR 09-SEP-1998; 98US-0099642.
PR 10-SEP-1998; 98US-0099741.
PR 10-SEP-1998; 98US-0099754.
PR 10-SEP-1998; 98US-0099763.
PR 10-SEP-1998; 98US-0099792.
PR 10-SEP-1998; 98US-0099808.
PR 10-SEP-1998; 98US-0099812.
PR 10-SEP-1998; 98US-0099815.
PR 15-SEP-1998; 98US-0099816.
PR 15-SEP-1998; 98US-0100385.
PR 15-SEP-1998; 98US-0100388.
PR 15-SEP-1998; 98US-0100390.
PR 16-SEP-1998; 98US-0100584.
PR 16-SEP-1998; 98US-0100627.
PR 16-SEP-1998; 98US-0100661.
PR 16-SEP-1998; 98US-0100662.
PR 16-SEP-1998; 98US-0100663.
PR 17-SEP-1998; 98US-0100683.
PR 17-SEP-1998; 98US-0100684.
PR 17-SEP-1998; 98US-0100710.
PR 17-SEP-1998; 98US-0100711.
PR 17-SEP-1998; 98US-0100919.
PR 17-SEP-1998; 98US-0100930.
PR 18-SEP-1998; 98US-0100848.
PR 18-SEP-1998; 98US-0100849.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101068.
PR 18-SEP-1998; 98US-0101071.
PR 22-SEP-1998; 98US-0101279.
PR 23-SEP-1998; 98US-0101471.
PR 23-SEP-1998; 98US-0101472.
PR 23-SEP-1998; 98US-0101474.
PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 23-SEP-1998; 98US-0101479.
PR 24-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101915.
PR 24-SEP-1998; 98US-0101916.

PR 29-SEP-1998; 98US-0102207.
PR 29-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102330.
PR 29-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102484.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102965.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103314.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103395.
PR 07-OCT-1998; 98US-0103396.
PR 08-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103633.
PR 08-OCT-1998; 98US-0103678.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103717.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.
PR 22-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105693.
PR 26-OCT-1998; 98US-0105694.
PR 26-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105881.
PR 27-OCT-1998; 98US-0105882.
PR 27-OCT-1998; 98US-0106062.
PR 28-OCT-1998; 98US-0106023.
PR 28-OCT-1998; 98US-0106029.
PR 28-OCT-1998; 98US-0106030.
PR 28-OCT-1998; 98US-0106032.
PR 28-OCT-1998; 98US-0106033.
PR 28-OCT-1998; 98US-0106178.
PR 29-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106384.
PR 29-OCT-1998; 98US-0106500.
PR 30-OCT-1998; 98US-0106500.
PR 30-OCT-1998; 98US-0106502.
PR 30-OCT-1998; 98US-0106502.
PR 03-NOV-1998; 98US-0106505.
PR 03-NOV-1998; 98US-0106519.
PR 03-NOV-1998; 98US-0106932.
PR 03-NOV-1998; 98US-0106934.
PR 10-NOV-1998; 98US-0107783.
PR 17-NOV-1998; 98US-0108775.
PR 17-NOV-1998; 98US-0108779.
PR 17-NOV-1998; 98US-0108787.
PR 17-NOV-1998; 98US-0108788.
PR 17-NOV-1998; 98US-0108801.
PR 17-NOV-1998; 98US-0108802.
PR 17-NOV-1998; 98US-0108806.
PR 17-NOV-1998; 98US-0108807.
PR 17-NOV-1998; 98US-0108867.
PR 17-NOV-1998; 98US-0108925.
PR 18-NOV-1998; 98US-0108848.
PR 18-NOV-1998; 98US-0108849.
PR 18-NOV-1998; 98US-0108850.
PR 18-NOV-1998; 98US-0108851.
PR 18-NOV-1998; 98US-0108852.
PR 18-NOV-1998; 98US-0108858.
PR 18-NOV-1998; 98US-0108904.
XX
PA (GETH) GENENTECH INC.

PS Claim 3; Fig 495; 813bp; English.

CC AA521244-AS31518 encode for novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of a peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.

SQ Sequence 771 BP; 169 A; 231 C; 195 G; 176 T; 0 other:

Query Match 86.6%; Score 661; DB 22; Length 771;
 Best Local Similarly 99.7%; Pred. No. 0;
 Matches 761; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTCACATGCAACCAACGAGCATGCTCCCGAGGCTGATGCTGACTGCTCTTGGCA 60
 DB 1 CTCACATGCAACCAACGAGCATGCTCCCGAGGCTGATGCTGACTGCTCTTGGCA 60
 QY 61 TTTTCTCATCTCCAGGCTCTCTGCTCACAGGAGCCCAAGTGGCCCAATGACTCTT 120
 DB 61 TTTTCTCATCTCCAGGCTCTCTGCTCACAGGAGCCCAAGTGGCCCAATGACTCTT 120
 QY 121 ACCGATGCTGTGGCAGCCACACAGAGATGTGGGACAGTCTACGACCCCTGACG 180
 DB 121 ACCGATGCTGTGGCAGCCACACAGAGATGTGGGACAGTCTACGACCCCTGACG 180
 QY 181 ACTGTTCTATGATGATGCGGCTGCTGGCCAGAGCCCAACAGCTGTGGAACCTCA 240
 DB 181 ACTGTTCTATGATGATGCGGCTGCTGGCCAGAGCCCAACAGCTGTGGAACCTCA 240
 QY 241 CCTTCAAGTCTGCTTGTGACAGTGTGCCCCCTGACCTTATGATGATGATTAACC 300
 DB 241 CCTTCAAGTCTGCTTGTGACAGTGTGCCCCCTGACCTTATGATGATGATTAACC 300
 QY 301 AGAATGCGACTACAGCCCGAGCTGGATGACAGGCTTTGTGCGAGTGTGACTAATGA 360
 DB 301 AGAATGCGACTACAGCCCGAGCTGGATGACAGGCTTTGTGCGAGTGTGACTAATGA 360
 QY 361 ACATGAGGGGAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 DB 361 ACATGAGGGGAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 421 GGTGTTACCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 DB 421 GGTGTTACCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 481 AACTGCCCATCTATCTGTGACTGTCTGAGGCCACCTGCGAGTGTGCTGAGGAGGC 540
 DB 481 AACTGCCCATCTATCTGTGACTGTCTGAGGCCACCTGCGAGTGTGCTGAGGAGGC 540
 QY 541 CCACAGGTCCCTTCTAGATTTGTGACACAGATGAGATGTGCTGATGAGGAGGCAG 600
 DB 541 CCACAGGTCCCTTCTAGATTTGTGACACAGATGAGATGTGCTGATGAGGAGGCAG 600
 QY 601 GGACTGTGAACCTCTCTGATGATGACCTTATGAGCCCAATCAACCCGGGACCAAGGC 660
 DB 601 GGACTGTGAACCTCTCTGATGATGACCTTATGAGCCCAATCAACCCGGGACCAAGGC 660

QY 661 TGGCTGGGGAACCTTCACCCCTCTGTGAGATTTTCCATCATCTCAAGTCTCTATC 720
 DB 661 TGGCTGGGGAACCTTCACCCCTCTGTGAGATTTTCCATCATCTCAAGTCTCTATC 720
 QY 721 CAGGAGCAAAAGCAGAGATCAATTAATTTATGACTTTATA 763
 DB 721 CAGGAGCAAAAGCAGAGATCAATTAATTTATGACTTTATA 763

RESULT 7
 AAC85962
 ID AAC85962 standard; cDNA; 771 BP.
 XX
 AC AAC85962;
 DT 22-AUG-2001 (first entry)
 XX
 DE Native sequence of PRO1274 cDNA, clone DNA64889-1541.
 XX
 KW PRO: PRO1081; PRO1274; PRO10272; proliferation; T-lymphocyte;
 KW PRO1199; PRO1556; PRO4401; PRO10268; inhibition; stimulation;
 KW infiltration; mononuclear cell; eosinophil; erythema multiforme;
 KW polymorphonuclear neutrophil; PMN; antibody; immune-related disorder;
 KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
 KW juvenile chronic arthritis; spondyloarthritis; systemic sclerosis;
 KW idiopathic inflammatory myopathy; Sjogren's syndrome; skin disease;
 KW systemic vasculitis; sarcoidosis; autoimmune haemolytic anaemia; asthma;
 KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus; allergy;
 KW immune-mediated renal disease; demyelination; central nervous system;
 KW peripheral nervous system; idiopathic demyelinating polyneuropathy;
 KW Guillain-Barre syndrome; hepatobiliary disease; eosinophilic pneumonia;
 KW granulomatous hepatitis; primary biliary cirrhosis; allergic rhinitis;
 KW inflammatory bowel disease; gluten-sensitive enteropathy; urticaria;
 KW Whipple's disease; idiopathic pulmonary fibrosis; contact dermatitis;
 KW psoriasis; atopic dermatitis; hypersensitivity pneumonitis;
 KW graft rejection; graft-versus-host disease; ss.

OS Homo sapiens.
 XX
 XX
 EH Key Location/Qualifiers
 FT CDS 24..356
 FT /*tag- a
 FT /product= "PRO1274 polypeptide"
 FT sig_peptide 24..95
 FT /*tag- b
 FT mat_peptide 96..353
 FT /*tag- c

W0200140465-A2.
 PD 07-JUN-2001.
 XX
 PF 10-NOV-2000; 2000MO-US30873.
 XX
 PR 30-NOV-1999; 99MO-US28313.
 PR 09-DEC-1999; 99US-0170262.
 PR 23-DEC-1999; 99US-0172059.
 PR 11-JAN-2000; 2000US-0175481.
 PR 20-JAN-2000; 2000US-0177118.
 PR 18-FEB-2000; 2000MO-US04342.
 PR 03-MAR-2000; 2000US-0187202.
 PR 30-MAY-2000; 2000MO-US14941.
 PR 05-JUN-2000; 2000US-0209832.
 PR 24-AUG-2000; 2000MO-US23328.
 XX
 XX (GENENTECH INC.)
 PA
 PI Fong S, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
 PI Tamas D, Watanabe CK, Wood WI, Zhang Z;
 XX WPI; 2001-381384/40.
 DR P-FSDB; ABA47292.

Sequence 771 BP; 169 A; 231 C; 195 G; 176 T; 0 other;

Query Match 86.6%; Score 661; DB 22; Length 771;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 761; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 CTCACATGCAACCAACCCAGACATGCTCCCGAGGCTGATGCTGCTTTGCA 60
   1 CTCACATGCAACCAACCCAGACATGCTCCCGAGGCTGATGCTGCTTTGCA 60
Db 1 CTCACATGCAACCAACCCAGACATGCTCCCGAGGCTGATGCTGCTTTGCA 60
QY 61 TTTTCTGATCTCCAGGCTCTCTGCTCACAGGAGCCCGAGTGGCCCAATGACTCTT 120
   61 TTTTCTGATCTCCAGGCTCTCTGCTCACAGGAGCCCGAGTGGCCCAATGACTCTT 120
Db 61 TTTTCTGATCTCCAGGCTCTCTGCTCACAGGAGCCCGAGTGGCCCAATGACTCTT 120
QY 121 ACCGATGCTGTGCGACACACAGAGATGTGGGAGCACTTTACGACCCCTGACG 180
   121 ACCGATGCTGTGCGACACACAGAGATGTGGGAGCACTTTACGACCCCTGACG 180
Db 121 ACCGATGCTGTGCGACACACAGAGATGTGGGAGCACTTTACGACCCCTGACG 180
QY 181 ACTGTTCTATGATGATGCGCGCTGCGCTGGCGAGGACCCAGACCTGTGGAATGCA 240
   181 ACTGTTCTATGATGATGCGCGCTGCGCTGGCGAGGACCCAGACCTGTGGAATGCA 240
Db 181 ACTGTTCTATGATGATGCGCGCTGCGCTGGCGAGGACCCAGACCTGTGGAATGCA 240
QY 241 CCTTCAGATCTGCTTTGAGCAGTGTGCCCTTGACCTTATGATGATGATTAAC 300
   241 CCTTCAGATCTGCTTTGAGCAGTGTGCCCTTGACCTTATGATGATGATTAAC 300
Db 241 CCTTCAGATCTGCTTTGAGCAGTGTGCCCTTGACCTTATGATGATGATTAAC 300
QY 301 AGAAGTGGGAGTACAGCCCGGAGCTGAGATGACAGGCTTTGTGCAAGTGTGAGTGA 360
   301 AGAAGTGGGAGTACAGCCCGGAGCTGAGATGACAGGCTTTGTGCAAGTGTGAGTGA 360
Db 301 AGAAGTGGGAGTACAGCCCGGAGCTGAGATGACAGGCTTTGTGCAAGTGTGAGTGA 360
QY 361 ACATCAGGGGAGAGATGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
   361 ACATCAGGGGAGAGATGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 361 ACATCAGGGGAGAGATGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 GGTGTTACCTGATGATGCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
   421 GGTGTTACCTGATGATGCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 421 GGTGTTACCTGATGATGCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 AACTGCGCAATTTATTTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTG 540
   481 AACTGCGCAATTTATTTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTG 540
Db 481 AACTGCGCAATTTATTTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTG 540
QY 541 CCACAGGTGCTGCTTATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 600
   541 CCACAGGTGCTGCTTATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 600
Db 541 CCACAGGTGCTGCTTATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 600
QY 601 GGACTCTGAACCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
   601 GGACTCTGAACCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db 601 GGACTCTGAACCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 TGGCTGGGAGACCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
   661 TGGCTGGGAGACCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 661 TGGCTGGGAGACCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 721 CAGGAGCAAGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 763
   721 CAGGAGCAAGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 763
Db 721 CAGGAGCAAGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 763

```

RESULT 9

AAA08111 standard; DNA; 914 BP.

AAA08111;

22-JUN-2000 (first entry)

Human bladder specific gene 22638 clone 819141H SEQ ID NO.1.

Human; bladder specific gene; BSG; bladder cancer; diagnosis;

KW detection; metastasis; cytostatic; ds.

OS Homo sapiens.

```

XX XX WO200012761-A1.
XX XX 09-MAR-2000.
XX XX 01-SEP-1999; 99WO-US20003.
XX XX 02-SEP-1998; 98US-0098825.
XX XX (DIAD-) DIADEXUS LLC.
XX XX Salceda S, Sun Y, Recipon H, Cafferkey R;
XX XX WPI; 2000-256658/22.
XX XX
XX XX Detecting, diagnosing metastasis, staging, monitoring, imaging and
XX XX treating bladder cancer, involves measuring bladder specific gene (BSG)
XX XX levels in cells or body fluids.
XX XX
XX XX Claim 6; Page 25-26; 31pp; English.
XX XX
XX XX A method has been developed for detecting, diagnosing metastasis,
XX XX staging and monitoring the onset of metastasis or a change in stage
XX XX of bladder cancer, by measuring bladder specific gene (BSG) levels in
XX XX cells, tissues or body fluids. An increase in BSG levels compared to
XX XX normal human control is associated with the presence or progression of
XX XX bladder cancer, a decrease is associated with cancer regression. The
XX XX method comprising measuring BSG is useful for detecting, diagnosing
XX XX metastasis, staging and monitoring the onset of metastasis or a change
XX XX in stage of bladder cancer. An antibody against BSG (I) labeled with
XX XX paramagnetic ions or an isotope is useful for imaging bladder cancer
XX XX and (I), especially conjugated to a cytotoxic agent is useful for
XX XX treating bladder cancer. The method is efficient in discriminating
XX XX between bladder cancer which has metastasised and the bladder cancer
XX XX which has not metastasised. The present sequence represents a BSG
XX XX sequence used in the exemplification of the present invention.

```

Sequence 914 BP; 193 A; 265 C; 237 G; 218 T; 1 other;

Query Match 86.6%; Score 661; DB 21; Length 914;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 761; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 CTCACATGCAACCAACCCAGACATGCTCCCGAGGCTGATGCTGCTTTGCA 60
   33 CTCACATGCAACCAACCCAGACATGCTCCCGAGGCTGATGCTGCTTTGCA 60
Db 33 CTCACATGCAACCAACCCAGACATGCTCCCGAGGCTGATGCTGCTTTGCA 60
QY 61 TTTTCTGATCTCCAGGCTCTCTGCTCACAGGAGCCCGAGTGGCCCATGACTCTT 120
   61 TTTTCTGATCTCCAGGCTCTCTGCTCACAGGAGCCCGAGTGGCCCATGACTCTT 120
Db 61 TTTTCTGATCTCCAGGCTCTCTGCTCACAGGAGCCCGAGTGGCCCATGACTCTT 120
QY 93 TTTTCTGATCTCCAGGCTCTCTGCTCACAGGAGCCCGAGTGGCCCATGACTCTT 152
   93 TTTTCTGATCTCCAGGCTCTCTGCTCACAGGAGCCCGAGTGGCCCATGACTCTT 152
Db 93 TTTTCTGATCTCCAGGCTCTCTGCTCACAGGAGCCCGAGTGGCCCATGACTCTT 152
QY 121 ACCGATGCTGTGCGACACACAGAGATGTGGGAGCAAGTGTACAGCCCTGACG 180
   121 ACCGATGCTGTGCGACACACAGAGATGTGGGAGCAAGTGTACAGCCCTGACG 180
Db 121 ACCGATGCTGTGCGACACACAGAGATGTGGGAGCAAGTGTACAGCCCTGACG 180
QY 153 ACCGATGCTGTGCGACACACAGAGATGTGGGAGCAAGTGTACAGCCCTGACG 212
   153 ACCGATGCTGTGCGACACACAGAGATGTGGGAGCAAGTGTACAGCCCTGACG 212
Db 153 ACCGATGCTGTGCGACACACAGAGATGTGGGAGCAAGTGTACAGCCCTGACG 212
QY 181 ACTGTTCTATGATGATGCGCGCTGCGCTGGCGAGGACCCAGAGCTGTGGAATGCA 240
   181 ACTGTTCTATGATGATGCGCGCTGCGCTGGCGAGGACCCAGAGCTGTGGAATGCA 240
Db 181 ACTGTTCTATGATGATGCGCGCTGCGCTGGCGAGGACCCAGAGCTGTGGAATGCA 240
QY 213 ACTGTTCTATGATGATGCGCGCTGCGCTGGCGAGGACCCAGAGCTGTGGAATGCA 272
   213 ACTGTTCTATGATGATGCGCGCTGCGCTGGCGAGGACCCAGAGCTGTGGAATGCA 272
Db 213 ACTGTTCTATGATGATGCGCGCTGCGCTGGCGAGGACCCAGAGCTGTGGAATGCA 272
QY 241 CCTTCAGATCTGCTTTGAGCAGTGTGCCCTTGACCTTATGATGATGATTAAC 300
   241 CCTTCAGATCTGCTTTGAGCAGTGTGCCCTTGACCTTATGATGATGATTAAC 300
Db 241 CCTTCAGATCTGCTTTGAGCAGTGTGCCCTTGACCTTATGATGATGATTAAC 300
QY 273 CCTTCAGATCTGCTTTGAGCAGTGTGCCCTTGACCTTATGATGATGATTAAC 332
   273 CCTTCAGATCTGCTTTGAGCAGTGTGCCCTTGACCTTATGATGATGATTAAC 332
Db 273 CCTTCAGATCTGCTTTGAGCAGTGTGCCCTTGACCTTATGATGATGATTAAC 332
QY 301 AGAAGTGGGAGTACAGCCCGGAGCTGAGATGACAGGCTTTGTGCAAGTGTGAGTGA 360
   301 AGAAGTGGGAGTACAGCCCGGAGCTGAGATGACAGGCTTTGTGCAAGTGTGAGTGA 360
Db 301 AGAAGTGGGAGTACAGCCCGGAGCTGAGATGACAGGCTTTGTGCAAGTGTGAGTGA 360
QY 333 AGAAGTGGGAGTACAGCCCGGAGCTGAGATGACAGGCTTTGTGCAAGTGTGAGTGA 392
   333 AGAAGTGGGAGTACAGCCCGGAGCTGAGATGACAGGCTTTGTGCAAGTGTGAGTGA 392
Db 333 AGAAGTGGGAGTACAGCCCGGAGCTGAGATGACAGGCTTTGTGCAAGTGTGAGTGA 392
QY 361 ACATCAGGGGAGAGATGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
   361 ACATCAGGGGAGAGATGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 361 ACATCAGGGGAGAGATGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 393 ACATCAGGGGAGAGATGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 452
   393 ACATCAGGGGAGAGATGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 452
Db 393 ACATCAGGGGAGAGATGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 452
QY 421 GGTGTTACCTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 480
   421 GGTGTTACCTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 480

```

Db 453 GGTGTACCTGAGATCTGGGATGCTGAGTGGCTGTTGGGGGCGAGAAACACACACTC 512
QY 481 AACTGCCACTTCAATTCGTGACCTGTGTAGAGCCACCTGGCCGTGCCCTGAGAGGC 540
Db 513 AACTGCCACTTCAATTCGTGACCTGTGTAGAGCCACCTGGCCGTGCCCTGAGAGGC 572
QY 541 CCAACAGTCCCTCTTCAATTCGTGAGACGATGAGATGCGTGTGTGTAGTGGGGCCGAG 600
Db 573 CCACAGGTCCTCTTCAATTCGTGAGACGATGAGATGCGTGTGTGTAGTGGGGCCGAG 632
QY 601 GGACTGTGAACCTCCGATGATGACCCATATGAGCCAAATCAACCCGACACCCCAAGGC 660
Db 633 GGACTGTGAACCTCCGATGATGACCCATATGAGCCAAATCAACCCGACACCCCAAGGC 692
QY 661 TGGCTGGGGAACCTTCACCTTCTGTGAGATTTCCATCATCTCAAGTTCCTTATC 720
Db 693 TGGCTGGGGAACCTTCACCTTCTGTGAGATTTCCATCATCTCAAGTTCCTTATC 752
QY 721 CAGAGCAAAACACAGATCATTAATTAATTTAGTACTTATA 763
Db 753 CAGAGCAAAACACAGATCATTAATTAATTTAGTACTTATA 795

RESULT 10
AAZ98158
ID AAZ98158 standard; cDNA; 748 BP.
XX
AC AAZ98158;
XX
DT 11-MAY-2000 (first entry)
XX
DE Human signal peptide containing protein HSP-50 cDNA SEQ ID NO:184.
XX
KW Human: signal peptide-containing protein; HSP; diagnosis; cancer;
KW inflammation; cardiovascular disease; anticancer; anti-inflammatory;
KW antimicrobial; nootropic; neuroprotective; cardiovascular; hepatotropic;
KW antidiabetic; gene therapy; cell proliferation; neurological disorder;
KW reproductive disorder; developmental disorder; arteriosclerosis;
KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;
KW asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;
KW Parkinson's disease; Huntington's diseases; ovulatory defect;
KW muscular dystrophy; ss.
XX
OS Homo sapiens.
XX
PN WO20000610-A2.
XX
PD 06-JAN-2000.
XX
PE 25-JUN-1999; 99MO-US14484.
XX
PR 26-JUN-1998; 98US-0090762.
PR 31-JUL-1998; 98US-0094983.
PR 01-OCT-1998; 98US-0102586.
PR 11-DEC-1998; 98US-0112129.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR;
PI Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hallman JL;
PI Bandman O;
XX
DR MPI; 2000-160673/14.
XX
PT P-PSDB; AAY87273.
XX
PT New human signal peptide-containing proteins useful in treatment,
PT prevention and diagnosis of e.g. cancer, inflammation and
XX
XX cardiovascular disease
XX
XX Claim 9; Page 281; 327pp; English.
XX
CC AAZ98109 to AAZ98342 encode AAY87224 to AAY87357 which represent the
CC human signal peptide-containing proteins HSP-1 to HSP-134. HSPs have

CC anticancer, anti-inflammatory, antimicrobial, nootropic, hepatotropic,
CC neuroprotective, cardiovascular, and antidiabetic activities, and can
CC be used in gene therapy. HSPs can be used to treat or prevent disorders
CC associated with decreased activity or function of HSP. Antagonists of
CC HSP are used to treat or prevent disorders associated with increased
CC activity or function of HSP. Such diseases include cell proliferation
CC (including cancer), inflammation, cardiovascular, neurological,
CC reproductive or developmental disorders, (e.g. arteriosclerosis,
CC cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,
CC asthma, Crohn's disease, microbial or other infections, congestive or
CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's
CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSP
CC nucleic acids can be used for the recombinant production of HSP, for
CC detecting HSP in standard hybridisation and amplification assays (for
CC diagnosis and monitoring), in gene therapy, as antisense,
CC triplex-forming or ribozyme therapeutics, for detecting related sequences
CC or genetic variations, and for chromosomal mapping. HSP are also used to
CC raise specific antibodies (Ab) and to screen for agonists and
CC antagonists (potential therapeutic agents). Ab are used to diagnose, or
CC monitor, HSP-related diseases (in usual immunoassays), as therapeutic
CC antagonists, in competitive drug screens, and for purification of HSP
CC from natural sources.
XX
SQ Sequence 748 BP; 159 A; 227 C; 193 G; 169 T; 0 other;
XX
Query Match 74.6%; Score 569; DB 21; Length 748;
Best Local Similarity 99.8%; Pred. No. 5e-275;
Matches 619; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6 CTGCACACCCAGAGCCATGCTCCCGAGAGCGTCAATGATGCTTGGCATTTTC 65
Db 1 CTGCACACCCAGAGCCATGCTCCCGAGAGCGTCAATGATGCTTGGCATTTTC 60
QY 66 TGCATCTCCAGGCTCTCTGCTCTACAGAGACCCCAATGGCCCATGACTCTTACTG 125
Db 61 TGCATCTCCAGGCTCTCTGCTCTACAGAGACCCCAATGGCCCATGACTCTTACTG 120
QY 126 ATGCTGGCCAGCCACAGAGATGAGGAGACAGTTTCAAGCCCTCCAGACTGT 185
Db 121 ATGCTGGCCAGCCACAGAGATGAGGAGACAGTTTCAAGCCCTCCAGACTGT 180
QY 186 TGCATGATGATGCGCTGTGCTGCTGGCCAGAGACCGAGAGTGGAGAACTGACCTTC 245
Db 181 TGCATGATGATGCGCTGTGCTGCTGGCCAGAGACCGAGAGTGGAGAACTGACCTTC 240
QY 246 AGAGTCTGCTTTAGAGCAGTGTGCTGCTGACCTTCAATGATGATTAACCAAGAC 305
Db 241 AGAGTCTGCTTTAGAGCAGTGTGCTGCTGACCTTCAATGATGATTAACCAAGAC 300
QY 306 TGCAGCTCAGCCCGAGACCTGAGATGACAGCTTGTGAGTGCACCTAATGGAACATC 365
Db 301 TGCAGCTCAGCCCGAGACCTGAGATGACAGCTTGTGAGTGCACCTAATGGAACATC 360
QY 366 AGGGGAACGATGACTCTGATTTCTCTTCTGCTGGGTGGCTGGAGAAAGCTGTGT 425
Db 361 AGGGGAACGATGACTCTGATTTCTCTTCTGCTGGGTGGCTGGAGAAAGCTGTGT 420
QY 426 TACCTGATGATGAGATGCTGAGTGGCTGTTGGGGCCAGAGAAACACACTCAACTG 485
Db 421 TACCTGATGATGAGATGCTGAGTGGCTGTTGGGGCCAGAGAAACACACTCAACTG 480
QY 486 CCCACTTCAATTCGTGACCTGTGAGGCGCCACCTGCGCTGCCGAGAGAGGCCACCA 545
Db 481 CCCACTTCAATTCGTGACCTGTGAGGCGCCACCTGCGCTGCCGAGAGAGGCCACCA 540
QY 546 GGTCCCTTCTAGAAATTCGTGAGACGATGATGCGTGTGTATGGGGCCAGAGACT 605
Db 541 GGTCCCTTCTAGAAATTCGTGAGACGATGATGCGTGTGTATGGGGCCAGAGACT 600
QY 606 CTGAACCTCTGATGACC 625
Db 601 CTGAACCTCTGATGACC 620

```
RESULT 11
AA01685/C
ID AAX01685 standard; DNA; 417 BP.
XX
XX AAX01685;
AC
XX
XX
XX 21-APR-1999 (first entry)
XX
XX Urinary tract tissue library BL172 gene specific cDNA clone g2178680.
XX
XX BL172; urinary tract; marker; cancer; recombinant; human; ss.
XX
XX Homo sapiens.
XX
XX W09855656-A1.
XX
XX 10-DEC-1998.
XX
XX 05-JUN-1998; 98MO-US11693.
XX
XX 05-JUN-1997; 97US-0869579.
XX
XX (ABBO ) ABBOT LAB.
XX
XX Billing-medel PA, Cohen M, Colpitts TL, Friedman PN,
XX Gordon J, Grandosen, Hodges SC, Klass MR, Kratochvil JD;
XX Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;
XX
XX WPI; 1999-045802/04.
XX
XX
XX PT New purified polynucleotide BL172 derivatives and encoded
XX PT polypeptides - useful in the identification of markers, which are
XX PT indicative of urinary tract diseases or conditions
XX
XX Claim 1; Page 89; 114pp; English.
XX
XX Sequences AAX01683-85 represent overlapping clones specific for urinary
XX tract tissue library BL172. The BL172 polynucleotides are used to detect
XX target BL172 polynucleotides and BL172 mRNA and the BL172 polypeptides
XX are used to detect BL172 antigens/antibodies in a test sample. The
XX identification of certain markers in these methods are indicative of the
XX presence of urinary tract disease, especially cancer. Additionally, the
XX polypeptides are used to detect antibodies that bind specifically to a
XX BL172 epitope derived from a BL172 polypeptide. Host cells containing a
XX recombinant expression vector comprising the BL172 polynucleotide
XX sequences are used to produce the polypeptides containing BL172
XX epitopes. The methods aid in diagnosis, detection, staging, monitoring,
XX or prognostication, in vivo imaging, prevention and treatment of diseases
XX or conditions associated with BL172, especially urinary tract cancer.
XX
XX Sequence 417 BP; 99 A; 108 C; 117 G; 93 T; 0 other;
XX
XX Query Match 48.0%; Score 366; DB 20; Length 417;
XX Best Local Similarity 99.8%; Pred. No. 2.8e-173;
XX Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 346 GTGTCAAGCTAATGAGCATCAGGAGGAGCATGACTCCTGATTCCTCTCTGGTGGC 405
XX |||||||
XX 417 GTGTCAAGCTAATGAGCATCAGGAGGAGCATGACTCCTGATTCCTCTCTGGTGGC 358
XX |||||||
XX 406 CTGGAGAAGAGGCTGCTTACTGTAGATCTGGAGTCTGAGTGGCTGTTGGGGCCA 465
XX |||||||
XX 357 CTGGAGAAGAGGCTGCTTACTGTAGATCTGGAGTCTGAGTGGCTGTTGGGGCCA 298
XX |||||||
XX 466 GAGAAGACACATCACTCACTCATCTGTGTGACCTGTGAGAGCCCAACCCCTCGC 525
XX |||||||
XX 297 GAGAAGACACATCACTCACTCATCTGTGTGACCTGTGAGAGCCCAACCCCTCGC 238
XX |||||||
XX 526 CTGACCTGAGAGAGCCACAGATCCCTTCTAGAAATTCCTGGAGAGCATGAGATCGTGT 585
XX |||||||
XX 237 CTGACCTGAGAGAGCCACAGATCCCTTCTAGAAATTCCTGGAGAGCATGAGATCGTGT 178
```

```
XX
XX 586 CTGATGGGGGCCACGAGACTGTGAACCTCTCTGATGACCCCTATGGCCAAATCAACCCG 645
XX |||||||
XX 177 CTGATGGGGGCCACGAGACTGTGAACCTCTCTGATGACCCCTATGGCCAAATCAACCCG 118
XX |||||||
XX 646 GCACGACCCCAAGGCTGGGTGGGGAACCTTCACCTCTGTGAGATTTTCATCATCTGC 705
XX |||||||
XX 117 GCACGACCCCAAGGCTGGGTGGGGAACCTTCACCTCTGTGAGATTTTCATCATCTGC 58
XX |||||||
XX 706 AAGTCTCTTCTATCCAGAGCAAGACAGATCATATAATTATGACTTAT 762
XX |||||||
XX 57 AAGTCTCTTCTATCCAGAGCAAGACAGATCATATAATTATGACTTAT 1
XX
XX
XX RESULT 12
XX AAX01684
XX ID AAX01684 standard; DNA; 236 BP.
XX
XX
XX AAX01684;
XX
XX 21-APR-1999 (first entry)
XX
XX Urinary tract tissue library BL172 gene specific cDNA clone 2624118.
XX
XX BL172; urinary tract; marker; cancer; recombinant; human; ss.
XX
XX Homo sapiens.
XX
XX W09855656-A1.
XX
XX 10-DEC-1998.
XX
XX 05-JUN-1998; 98MO-US11693.
XX
XX 05-JUN-1997; 97US-0869579.
XX
XX (ABBO ) ABBOT LAB.
XX
XX Billing-medel PA, Cohen M, Colpitts TL, Friedman PN;
XX Gordon J, Grandosen, Hodges SC, Klass MR, Kratochvil JD;
XX Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;
XX
XX WPI; 1999-045802/04.
XX
XX
XX PT New purified polynucleotide BL172 derivatives and encoded
XX PT polypeptides - useful in the identification of markers, which are
XX PT indicative of urinary tract diseases or conditions
XX
XX Claim 1; Page 89; 114pp; English.
XX
XX Sequences AAX01683-85 represent overlapping clones specific for urinary
XX tract tissue library BL172. The BL172 polynucleotides are used to detect
XX target BL172 polynucleotides and BL172 mRNA and the BL172 polypeptides
XX are used to detect BL172 antigens/antibodies in a test sample. The
XX identification of certain markers in these methods are indicative of the
XX presence of urinary tract disease, especially cancer. Additionally, the
XX polypeptides are used to detect antibodies that bind specifically to a
XX BL172 epitope derived from a BL172 polypeptide. Host cells containing a
XX recombinant expression vector comprising the BL172 polynucleotide
XX sequences are used to produce the polypeptides containing BL172
XX epitopes. The methods aid in diagnosis, detection, staging, monitoring,
XX or prognostication, in vivo imaging, prevention and treatment of diseases
XX or conditions associated with BL172, especially urinary tract cancer.
XX
XX Sequence 236 BP; 49 A; 69 C; 62 G; 56 T; 0 other;
XX
XX Query Match 30.9%; Score 236; DB 20; Length 236;
XX Best Local Similarity 100.0%; Pred. No. 4e-108;
XX Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 162 TTCTAGACCCCTCGACGACTGTGCTATGATGATCCCTGCGCCCTTGCCAGGACC 221
XX |||||||
XX 1 TTCTAGACCCCTCGACGACTGTGCTATGATGATCCCTGCGCCCTTGCCAGGACC 60
```

QY 222 CAGAGTGTGGAACAGCTCAGAGTCTGCTTTGAGCAGTCTGCCCCCTGAGCTTC 281
 |||||||
 DB 61 CAGACGTGTGGAACAGCTCAGAGTCTGCTTTGAGCAGTCTGCCCCCTGAGCTTC 120
 |||||||
 QY 282 ATGATGAGAGCTGATTAACAGCACTGCGAGCTGACCGGACCTGGATGACAGGCTTGT 341
 |||||||
 DB 121 ATGTTGAAGCTGATTAACAGCACTGCGAGCTGACCGGACCTGGATGACAGGCTTGT 180
 |||||||
 QY 342 CGCAGTGTGAGCTAATGGAACATCAGGGGAGAGATGACTCTGATTCCTCTCTCT 397
 |||||||
 DB 181 CGCAGTGTGAGCTAATGGAACATCAGGGGAGAGATGACTCTGATTCCTCTCTCT 236
 |||||||

RESULT 13
 AAX01683
 ID AAX01683 standard; DNA; 196 BP.
 XX
 AC AAX01683;
 XX
 DT 21-APR-1999 (first entry)
 XX
 DE Urinary tract tissue library BL172 gene specific cDNA clone 1554838.
 XX
 KM BL172; urinary tract; marker; cancer; recombinant; human; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9855656-A1.
 XX
 PD 10-DEC-1998.
 XX
 PF 05-JUN-1998; 98MO-US11693.
 XX
 PR 05-JUN-1997; 97US-0869579.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 PI Billing-medell PA, Cohen M, Colpitts TL, Friedman PN;
 PI Gordon J, Granadosen, Hodges SC, Klass MR, Kratochvil JD;
 PI Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;
 XX
 DR WPI; 1999-045802/04.
 XX
 PT New purified polynucleotide BL172 derivatives and encoded
 PT polypeptides - useful in the identification of markers, which are
 PT indicative of urinary tract diseases or conditions
 XX
 PS Claim 1; Page 89; 114pp; English.
 XX

CC Sequences AAX01683-85 represent overlapping clones specific for urinary
 CC tract tissue library BL172. The BL172 polynucleotides are used to detect
 CC target BL172 polynucleotides and BL172 mRNA and the BL172 polypeptides
 CC are used to detect BL172 antigens/antibodies in a test sample. The
 CC identification of certain markers in these methods are indicative of the
 CC presence of urinary tract disease, especially cancer. Additionally, the
 CC polypeptides are used to detect antibodies that bind specifically to a
 CC BL172 epitope derived from a BL172 polypeptide. Host cells containing a
 CC recombinant expression vector comprising the BL172 polynucleotide
 CC sequences are used to produce the polypeptides containing BL172
 CC epitopes. The methods aid in diagnosis, detection, staging, monitoring,
 CC prognosis, in vivo imaging, prevention and treatment of diseases
 CC or conditions associated with BL172, especially urinary tract cancer.
 CC
 XX Sequence 196 BP; 38 A; 70 C; 44 G; 44 T; 0 other;
 SQ

Query Match 25.7%; Score 196; DB 20; Length 196;
 Best Local Similarity 100.0%; Pred. No. 4.4e-88;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTCGACTGCAGACACCCAGAGGCTGCTCCGAGAGCTGATCGTGTCTTTGGCA 60
 |||||||
 DB 1 CTCGACTGCAGACACCCAGAGGCTGCTCCGAGAGCTGATCGTGTCTTTGGCA 60
 |||||||

QY 61 TTTTCTCAGCTCCAGGCTCCTCTGCTCAGACGAGGCCCCAGTGGCCCCATGACTCTT 120
 |||||||
 DB 61 TTTTCTCAGCTCCAGGCTCCTCTGCTCAGACGAGGCCCCAGTGGCCCCATGACTCTT 120
 |||||||
 QY 121 ACCGATGCTGTGGCCACACAGAGATGTGGGGACAGTTTACAGCCCCCTGAGC 180
 |||||||
 DB 121 ACCGATGCTGTGGCCACACAGAGATGTGGGGACAGTTTACAGCCCCCTGAGC 180
 |||||||
 QY 181 ACTGTGCTATGATGA 196
 |||||||
 DB 181 ACTGTGCTATGATGA 196
 |||||||

RESULT 14
 AAS92056
 ID AAS92056 standard; cDNA; 1338 BP.
 XX
 AC AAS92056;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #27860.
 XX
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001MO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 XX
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 XX
 DR P-PSDB; ABG27869.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 1; SEQ ID NO 27860; 103pp; English.
 XX

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 1338 BP; 311 A; 350 C; 348 G; 329 T; 0 other;

Query Match 3.0%; Score 23; DB 23; Length 1338;

Best Local Similarity 100.0%; Pred. No. 0.23;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AACTGCACCTTCAGAGTCTGCTT 256

Db 1266 AACTGCACCTTCAGAGTCTGCTT 1268

RESULT 15

AAAX01690
ID AAX01690 standard; DNA; 22 BP.

XX AAX01690;

AC 21-APR-1999 (first entry)

XX BL172 gene amplifying sense primer.

DE BL172; urinary tract; marker; cancer; recombinant; human; PCR primer; ss.

XX Synthetic.

OS Homo sapiens.

XX W09855656-A1.

XX 10-DEC-1998.

XX 05-JUN-1998; 98MO-US11693.

XX 05-JUN-1997; 97US-0869579.

XX (ABBO) ABBOTT LAB.

XX Billing-medel PA, Cohen M, Colpitts TL, Friedman PN;

XX Gordon J, Granadosen, Hodges SC, Klass MR, Kratochvil JD;

XX Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;

XX WPI; 1999-045802/04.

XX New purified polynucleotide BL172 derivatives and encoded

XX polypeptides - useful in the identification of markers, which are

XX indicative of urinary tract diseases or conditions

XX Example 8; Page 64; 114pp; English.

XX The invention provides polynucleotide sequences from urinary tract

XX tissue library BL172. The BL172 polynucleotides (AAAX01683-87) are used to

XX detect target BL172 polynucleotides and BL172 mRNA and the BL172

XX polypeptides are used to detect BL172 antigens/antibodies in a test

XX sample. The identification of certain markers in these methods are

XX indicative of the presence of urinary tract disease, especially cancer.

XX Additionally, the polypeptides are used to detect antibodies that bind

XX specifically to a BL172 epitope derived from a BL172 polypeptide. Host

XX cells containing a recombinant expression vector comprising the BL172

XX polynucleotide sequences are used to produce the polypeptides containing

XX BL172 epitopes. The methods aid in diagnosis, detection, staging,

XX monitoring, prognosis, in vivo imaging, prevention and treatment of

XX diseases or conditions associated with BL172, especially urinary tract

XX cancer. Sequences AAX01690-91 represent PCR primers used for the

XX amplification of BL172 cDNA.

XX SQ Sequence 22 BP; 3 A; 5 C; 6 G; 8 T; 0 other;

XX Query Match 2.9%; Score 22; DB 20; Length 22;

XX Best Local Similarity 100.0%; Pred. No. 0.71;

XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 AGGCTGACATGCTAGCTGCTTT 56

Db 1 AGGCTGACATGCTAGCTGCTTT 22

Search completed: November 3, 2002, 01:06:40
Job time : 214 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2002, 01:06:51 ; Search time 1689 Seconds
(without alignments)
6097.204 Million cell updates/sec

Title: US-09-092-297-5

Perfect score: 763
Sequence: 1 CTCCTACTGCACCAACCACCA.....ATAATTGACTTATA 763

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size: 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST.*
1: em_estdb:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_lin:*
15: em_gss_pln:*
16: em_gss_vrl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	410	53.7	828	10	BG775668
2	379	49.7	412	9	AA456370
3	371	48.6	533	9	AA991605
4	366	48.0	417	9	AA455904
5	328	43.0	604	10	BG680336
6	318	41.7	729	10	BG775059
7	263	34.5	354	9	AM265494
8	258	33.8	475	9	AA195677
9	235	30.8	253	10	BF911380
10	193	25.3	308	9	AM265432
11	163	21.4	201	10	BF911379
12	158	20.7	245	10	BF913866
13	116	15.2	245	10	BF914311
14	74	9.7	275	9	AI572169
15	60	7.9	103	10	BG951874
16	48	6.3	130	9	AM795497
17	26	3.4	429	9	AM424984

C 18	23	3.0	1169	10	BI490599
C 19	21	2.8	428	9	BF796894
C 20	21	2.8	654	10	BE291082
C 21	20	2.6	271	9	BE230495
C 22	20	2.6	323	9	BE098543
C 23	20	2.6	398	12	A2049839
C 24	20	2.6	418	10	BG880839
C 25	20	2.6	433	9	AU163926
C 26	20	2.6	458	10	C20067
C 27	20	2.6	484	10	BI336276
C 28	20	2.6	547	10	BF078123
C 29	20	2.6	679	9	BE230496
C 30	20	2.6	957	10	BI768819
C 31	19	2.5	280	9	BB256375
C 32	19	2.5	281	9	BB838433
C 33	19	2.5	328	10	BC091023
C 34	19	2.5	344	10	BF448934
C 35	19	2.5	360	9	AV188034
C 36	19	2.5	360	10	D68665
C 37	19	2.5	364	12	A2098866
C 38	19	2.5	374	10	BF659363
C 39	19	2.5	376	9	AI955654
C 40	19	2.5	415	9	AM205354
C 41	19	2.5	437	10	H83681
C 42	19	2.5	482	9	AM677348
C 43	19	2.5	483	10	BG394735
C 44	19	2.5	505	12	BH225730
C 45	19	2.5	509	12	AZ953152

ALIGNMENTS

RESULT 1
BG775668/c 828 bp mRNA linear EST 15-MAY-2001
60265002371 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4761029 3',
LOCUS
DEFINITION
mRNA sequence.
ACCESSION
BG775668
VERSION
BG775668.1 GI:14045965
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 828)
NIH-MGC http://mgs.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaps-femail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM612 row: n column: 06
High quality sequence start: 16
High quality sequence stop: 806.
Location/Qualifiers

FEATURES
source
1. 828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4761029"
/clone_id="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pORF7; Site_1: xhoi;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(c). Library constructed by

Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 190 a 214 c 234 g 190 t

ORIGIN

Query Match 53.7%; Score 410; DB 10; Length 828;
Best Local Similarity 99.5%; Pred. No. 4.9e-197;
Matches 630; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 61 TTTTCGATCTCCAGGCTCTCTGCTCAGCAGAGCCGCCAGTGGCCCTGACTCTT 120
|||||
Db 658 TTTTCTCATCTCCAGGCTCTCTGCTCAGCAGAGCCGCCAGTGGCCCTGACTCTT 599
QY 121 ACCTGATGCTGTGCCAGCAGACAGAGATGTGGGACAAAGTTCTAGACCCCTGCAGC 180
|||||
Db 598 ACCTGATGCTGTGCCAGCAGACAGAGATGTGGGACAAAGTTCTAGACCCCTGCAGC 539
QY 181 ACTGTGCTATGATGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
|||||
Db 538 ACTGTGCTATGATGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
QY 241 CCTTCAGAGTGTCTTTGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
|||||
Db 478 CCTTCAGAGTGTCTTTGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 301 AGAAGTGCAGTGCAGCCGCGAGCTGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 360
|||||
Db 419 AGAAGTGCAGTGCAGCCGCGAGCTGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 360
QY 361 ACATCAGGGGAGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
|||||
Db 359 ACATCAGGGGAGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 421 GGTGTACCTGAGATGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGT 480
|||||
Db 299 GGTGTACCTGAGATGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGT 240
QY 481 AACTGCCACATCTATCTGTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGT 540
|||||
Db 239 AACTGCCACATCTATCTGTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGT 180
QY 541 CCACAGTCTCCCTTCTAGAAATCTGAGACAGTGTGCTGAGTGTGCTGAGTGTGCTGAGT 600
|||||
Db 179 CCACAGTCTCCCTTCTAGAAATCTGAGACAGTGTGCTGAGTGTGCTGAGTGTGCTGAGT 120
QY 601 GGAAGTGCAGTGCAGCCGCGAGCTGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 660
|||||
Db 119 GGAAGTGCAGTGCAGCCGCGAGCTGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 60
QY 661 TGGCTGGGGAACCTTCACCTTCTGTGAGATT 693
|||||
Db 59 TGGCTGGGGAACCTTCACCTTCTGTGAGATT 27

RESULT 2
AA456370 412 bp mRNA linear EST 06-JUN-1997
LOCUS aa14602.r1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:813242
DEFINITION 5', mRNA sequence.

ACCESSION AA456370
VERSION AA456370.1 GI:2178946
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo. Hillier, L., Allen, M., Bowles, L., Dubuque, J., Geisels, G., Jost, S., Kucaba, T., Lacy, N., Le, N., Lennon, G., Matlin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theisling, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.

TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997).
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -28ml3 rev2 RT from Amersham
High quality sequence stop: 398.
location/Qualifiers
1. 412

FEATURES
source
/organism="Homo sapiens"
/db_xref="GDB:6044152"
/db_xref="taxon:9606"
/clone="IMAGE:813242"
/clone_1bp="Soares_NhHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"

/note="Organ: mixed (see below); Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NBH, pregnant uterus NBHPU, and fetal heart NBH199) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT 88 a 116 c 119 g 89 t

ORIGIN

Query Match 49.7%; Score 379; DB 9; Length 412;
Best Local Similarity 100.0%; Pred. No. 2.4e-181;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 CCCATATCTCTTACTGATGCTGTGCGACCCACACAGAGATGGGGACAAAGTTCTA 166
|||||
Db 34 CCCATATCTCTTACTGATGCTGTGCGACCCACACAGAGATGGGGACAAAGTTCTA 93
QY 167 CGACCCCTGAGACGCTGTGATGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 226
|||||
Db 94 CGACCCCTGAGACGCTGTGATGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 153
QY 227 GTGTGAAAGTGCAGCTTTCAGAGTGTGCTGTGAGTGTGCTGTGAGTGTGCTGTGAGT 286
|||||
Db 154 GTGTGAAAGTGCAGCTTTCAGAGTGTGCTGTGAGTGTGCTGTGAGTGTGCTGTGAGT 213
QY 287 GAAGCTATTAACACAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 346
|||||
Db 214 GAAGCTATTAACACAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 273
QY 347 TGTACGCTAATGGAATCATCAGGGGAGAGTATCTGTGATCTCTCCGCTGCTGCTGCTGCT 406
|||||
Db 274 TGTACGCTAATGGAATCATCAGGGGAGAGTATCTGTGATCTCTCCGCTGCTGCTGCTGCT 333
QY 407 TGGAGAAAGAGGCTGTGTTTACTGAGATCTGGGATCTGAGTGTGCTGTTTGGGGCCAG 466
|||||
Db 334 TGGAGAAAGAGGCTGTGTTTACTGAGATCTGGGATCTGAGTGTGCTGTTTGGGGCCAG 393
QY 467 AGAAACACACATCACTG 485
|||||
Db 394 AGAAACACACATCACTG 412

RESULT 3
AW991605 533 bp mRNA linear EST 05-JUN-2000
LOCUS AW991605
DEFINITION RCI-BN0005-030500-015-g07 BN0005 Homo sapiens cDNA, mRNA sequence.

ACCESSION AM91605
 VERSION AM91605.1 GI:8251660
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 533)
 AUTHORS Dias Neto,E., Garcia Coria,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2=RCI-BN0005-030>)
 500-015-9076t3-2000-05-036t4-1)
 Seq primer: puc 18 forward
 High quality sequence start: 16
 High quality sequence stop: 533.
 Location/Qualifiers
 1. 533
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="BN0005"
 /dev_stage="Adult"
 /note="Organ: Breast,normal; Vector: puc18; Site,1: SmaI; Site,2: SmaI; A mini-library was made by cloning products derived from ORESSES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc-18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 116 a 156 c 150 g 111 t
 ORIGIN
 Query Match 48.6%; Score 371; DB 9; Length 533;
 Best Local Similarity 99.8%; Pred. No. 2.9e-177;
 Matches 421; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 204 GTGCCCTTGCCAGGACCCAGAGCTGTGGAACCTGACCTTGAGTCTGCTTGAGCAG 253
 Db 78 GTGCCCTTGCCAGGACCCAGAGCTGTGGAACCTGACCTTGAGTCTGCTTGAGCAG 137
 Oy 264 TCGTGGCCCTTGACCTTCTGAGTGTGATTAACCAAGCTGCGACGCGCGGAGC 323
 Db 138 TCGTGGCCCTTGACCTTCTGAGTGTGATTAACCAAGCTGCGACGCGCGGAGC 197
 Oy 324 TCGGATGACAGCTTGTGTGACAGTGTACCTAATGAGACATCAGGGGAAAGATGACTCC 383
 Db 198 TCGGATGACAGCTTGTGTGACAGTGTACCTAATGAGACATCAGGGGAAAGATGACTCC 257
 Oy 384 GGATTCCTCTCTCTGGGTGGGCTGGAAGAAAGAGCTGTGTTACTGTGAGTGGGAGT 443
 Db 258 GGATTCCTCTCTCTGGGTGGGCTGGAAGAAAGAGCTGTGTTACTGTGAGTGGGAGT 317
 Oy 444 CTGAGTGGCTGTTTGGGGCCAGAGAAACACACACTCACTGCCCTCATTTCTGTGAC 503
 Db 318 CTGAGTGGCTGTTTGGGGCCAGAGAAACACACACTCACTGCCCTCATTTCTGTGAC 377

Oy 504 CTGTGTGAGGCCACCTGCGCTGCGCTGAGAGAGCCCAAGAGTCCCTTGTAGATTTC 563
 Db 378 CTGTGTGAGGCCACCTGCGCTGCGCTGAGAGAGCCCAAGAGTCCCTTGTAGATTTC 437
 Oy 564 TGGACAGCATGAGATGCTGTGCTGATGAGGGGCCAGGAGCTGTGACCTGTGATGAC 623
 Db 438 TGGACAGCATGAGATGCTGTGCTGATGAGGGGCCAGGAGCTGTGACCTGTGATGAC 497
 Oy 624 CC 625
 Db 498 CC 499
 RESULT 4
 AA455904/c 417 bp mRNA linear EST 06-JUN-1997
 LOCUS aa14602.s1 Soares_NHMPU_S1 Homo sapiens cDNA clone IMAGE:813242
 DEFINITION 3' mRNA sequence.
 ACCESSION AA455904
 VERSION AA455904
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 417)
 AUTHORS Hallier,L., Allen,M., Bowles,L., Dubnue,T., Gelsel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Scheilenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R.
 Washu-Merck EST Project 1997
 TITLE Unpublished (1997)
 JOURNAL Contact: Wilson RK
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LMD; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -41m13 fwd. ET from Amersham
 High quality sequence stop: 400.
 Location/Qualifiers
 1. 417
 /organism="Homo sapiens"
 /db_xref="GDB:6044152"
 /db_xref="taxon:9606"
 /clone="IMAGE:813242"
 /tissue_type="Soares_NHMPU_S1"
 /tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
 /lab_host="DH10B"
 /note="Organ: mixed (see below); Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site,1: Not I; Site,2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NBH, pregnant uterus NBHPU, and fetal heart NBH15W) were mixed, and ss circles were made in vitro. Following NRP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."
 BASE COUNT 99 a 108 c 117 g 93 t
 ORIGIN
 Query Match 48.0%; Score 366; DB 9; Length 417;
 Best Local Similarity 99.8%; Pred. No. 9.4e-175;
 Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 346 GTGTGAGCTAATGGAACATCAGGGGGAAGATGACTGATTCCTCTGAGTGGGAC 405
 Db 417 GTGTGAGCTAATGGAACATCAGGGGGAAGATGACTGATTCCTCTGAGTGGGAC 358

```

QY 406 CTGAGAAAGAGGCTGTGTACCTGAGATCTGGATGCTAGTGCTGTTGGGGCCA 465
Db 357 CTGAGAAAGAGGCTGTGTACCTGAGATCTGGATGCTAGTGCTGTTGGGGCCA 298
QY 466 GAGAAACACACACTCACTGCTCCACTTCTGCTGTGACCTGTGAGGCCACCTGCG 525
Db 297 GAGAAACACACACTCACTGCTCCACTTCTGCTGTGACCTGTGAGGCCACCTGCG 238
QY 526 CTGCGCTGAGAGAGGCCACAGTCCCTTCTAGATCTGAGACGATGATGCGGTG 585
Db 237 CTGCGCTGAGAGAGGCCACAGTCCCTTCTAGATCTGAGACGATGATGCGGTG 178
QY 586 CTGATGGGGGCCACAGGACTCTGAACCTCTGATGACCCGTATGGCCAAATCAACCC 645
Db 177 CTGATGGGGGCCACAGGACTCTGAACCTCTGATGACCCGTATGGCCAAATCAACCC 118
QY 646 GCACACACCCAGGCTGTGGGGAAACCTTCAACCTTCTGATGATTTCCATCATCTC 705
Db 117 GCACACACCCAGGCTGTGGGGAAACCTTCAACCTTCTGATGATTTCCATCATCTC 58
QY 706 AAGTCTCTTCTATCCAGAGCAAGACAGATCAATAAATTTATGACTTTAT 762
Db 57 AAGTCTCTTCTATCCAGAGCAAGACAGATCAATAAATTTATGACTTTAT 1

RESULT 5
LOCUS BG680336 604 bp mRNA linear EST 01-MAY-2001
DEFINITION 602629230F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4754071 5',
ACCESSION BG680336
VERSION BG680336.1 GI:13911733
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 604)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM10614 row: 1 column: 08
High quality sequence stop: 418.
Location/Qualifiers
1. 604
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4754071"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (TI phage-resistant)"
/Note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP library."
BASE COUNT 170 a 171 c 140 g 118 t 5 others
ORIGIN
Query Match 43.0%; Score 328; DB 10; Length 604;
Best Local Similarity 99.7%; Pred. No. 2e-155;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 329 TGACAGCCTTGTGCGAGTGTACGCTAATGGAACATCAGGGGAACGATGACTCTGAT 388

```

```

Db 14 TGACAGCCTTGTGCGAGTGTACGCTAATGGAACATCAGGGGAACGATGACTCTGAT 73
QY 389 CTCCTTCCTGGTGGGGCTGAGAAAGAGCTGTGTACTGAGATCTGGATGCTGAG 448
Db 74 CTCCTTCCTGGTGGGGCTGAGAAAGAGCTGTGTACTGAGATCTGGATGCTGAG 133
QY 449 TGAGTGTGGGGGCCAGAGAAACACACTCACTGCTCCACTTCTGCTGTGACCTGTC 508
Db 134 TGAGTGTGGGGGCCAGAGAAACACACTCACTGCTCCACTTCTGCTGTGACCTGTC 193
QY 509 TGAGGCCACCTGCGCTGCTCCCTGAGAGAGGCCACAGTCCCTTCTAGATTTCTGAC 568
Db 194 TGAGGCCACCTGCGCTGCTCCCTGAGAGAGGCCACAGTCCCTTCTAGATTTCTGAC 253
QY 569 AGCATGATGCGGTGCTATGGGGGCCAGGAGCTCTAACCCTCATGACCCCTA 628
Db 254 AGCATGATGCGGTGCTATGGGGGCCAGGAGCTCTAACCCTCATGACCCCTA 313
QY 629 TGCCCAACATCAACCCGCGACACACCCCAAGGCTGGTGGGAAACCTTCAACCTTCTGTG 688
Db 314 TGCCCAACATCAACCCGCGACACACCCCAAGGCTGGTGGGAAACCTTCAACCTTCTGTG 373
QY 689 AGATTTTCCATCATCTCAA 707
Db 374 AGATTTTCCATCATCTCAA 392

RESULT 6
LOCUS BG775059 729 bp mRNA linear EST 15-MAY-2001
DEFINITION 602650023F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4761029 5',
ACCESSION BG775059
VERSION BG775059.1 GI:14045376
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 729)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: DCTD/DP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM1612 row: n column: 06
High quality sequence stop: 726.
Location/Qualifiers
1. 729
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4761029"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: prostate; Vector: pOMB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC library."
BASE COUNT 169 a 212 c 181 g 166 t 1 others
ORIGIN

```

Query Match 41.7%; Score 318; DB 10; Length 729;
 Best Local Similarity 99.8%; Pred. No. 2.5e-150;
 Matches 438; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 61 TTTTCTGCATCTCCAGGCTCTCTGCTGCTACACGAGCCCGCCAGTGGCCCATGACTCCTT 120
 DB 2 TTTTCTGCATCTCCAGGCTCTCTGCTGCTACACGAGCCCGCCAGTGGCCCATGACTCCTT 61
 QY 121 ACCGATGCTGTGCGCAGCCACACAGAGATGTGGGAGCAATTTACGACCCCTGCAGC 180
 DB 62 ACCGATGCTGTGCGCAGCCACACAGAGATGTGGGAGCAATTTACGACCCCTGCAGC 121
 QY 181 ACCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 239
 DB 122 ACCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 181
 QY 240 ACCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 299
 DB 182 ACCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 241
 QY 300 CAGACTGCGACTCAGCCCGGACCTCGATGACAGAGCTTTGTCGAGTGTGAGCTAATG 359
 DB 242 CAGACTGCGACTCAGCCCGGACCTCGATGACAGAGCTTTGTCGAGTGTGAGCTAATG 301
 QY 360 AATATGAGGAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 419
 DB 302 AATATGAGGAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 361
 QY 420 TGGTGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 479
 DB 362 TGGTGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 421
 QY 480 CAATGCGCCACTTCATCT 498
 DB 422 CAATGCGCCACTTCATCT 440

RESULT 7 354 bp mRNA linear EST 28-DEC-1999
 LOCUS xq70c11.x1 NCI_CGAP_HN7 Homo sapiens CDNA clone IMAGE:2755988 3'
 DEFINITION mRNA sequence.

ACCESSION AW265494
 VERSION AW265494.1 GI:6642310
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 354)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL COMMENT Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 EMAIL: cga@bbs-rtmail.nih.gov
 TISSUE Procurement: Edward Shillitoe Ph.D., Silvio Gutkind Ph.D.,
 Chidchanok Leethanakul D.D.S., Michael Emmert-Buck M.D. Ph.D.,
 CDNA Library Preparation: David B. Krizman, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/dbp/image/image.html

FEATURES
 SOURCE
 1. 354
 Location/Qualifiers
 Seq primer: -400p from g1cgc.

possible reversed clone: polyT not found
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2755988"

/clone.lib="NCI_CGAP_HN7"
 /tissue.type="normal squamous epithelium, floor of mouth"
 /lab_host="DH10B"
 /note="Vector: pAMP10; CDNA made by oligo-dt priming.
 Non-directionally cloned into the UDG sites of pAMP10.
 Size-selected on agarose gel, average insert size 500 bp.
 Primary library; non-amplified. CDNA Library
 Preparation: David B. Krizman, Ph.D (NCI). Reference:
 Krizman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT 99 a 102 c 78 g 75 t
 Query Match 34.5%; Score 263; DB 9; Length 354;
 Best Local Similarity 99.7%; Pred. No. 1.8e-122;
 Matches 313; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 450 GCGTGTGGGGGCGCAGAAACACACACTCACTGCCACTTATCTGTGACTGTCT 509
 DB 9 GCGTGTGGGGGCGCAGAAACACACACTCACTGCCACTTATCTGTGACTGTCT 68
 QY 510 GAGGCCACCTGCGCGCTGAGGAGGCCACAGTCCCTTCTGATTTCTGGACA 569
 DB 69 GAGGCCACCTGCGCGCTGAGGAGGCCACAGTCCCTTCTGATTTCTGGACA 128
 QY 570 GCATGATGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 629
 DB 129 GCATGATGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 188
 QY 630 GCGCAATCAACCCCGGACACCCCAAGGCTGGGGAACCTTCACCTCTGTGA 689
 DB 189 GCGCAATCAACCCCGGACACCCCAAGGCTGGGGAACCTTCACCTCTGTGA 248
 QY 690 GATTTCCATCATCTCAAGTCTCTCTATCCAGAGCAAGACAGATCATATAAT 749
 DB 249 GATTTCCATCATCTCAAGTCTCTCTATCCAGAGCAAGACAGATCATATAAT 308
 QY 750 TTATGTACTTTATA 763
 DB 309 TTATGTACTTTATA 322

RESULT 8 475 bp mRNA linear EST 19-MAY-1997
 LOCUS zr32b04.s1 Soares_NHMPU_S1 Homo sapiens CDNA clone IMAGE:665143
 DEFINITION 3', mRNA sequence.

ACCESSION AA195677
 VERSION AA195677.1 GI:1785355
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 475)
 AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,
 Chisoe, S., Dietrich, N., Dubuque, T., Favallo, A., Gish, W., Hawkins,
 M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore,
 Schellenberg, K., Soares, M.B., Tan, F., Thierly-Meg, V., Treva, S.,
 Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marris, M.
 Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)
 JOURNAL MEDLINE 97044478
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 857 Std Error: 0.00
 High quality sequence stop: 308.

FEATURES
source

Location/Qualifiers
1. 475
/organism="Homo sapiens"
/db_xref="GDB:5427175"
/db_xref="taxon:9606"
/clone_image="65143"
/clone_lid="Soares.NhMPu.S1"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NBHM, pregnant uterus NbMPu, and fetal heart NbH15M) were mixed, and ss cicles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT 109 a 122 c 134 g 108 t 2 others

ORIGIN

Query Match 33.8%; Score 258; DB 9; Length 475;
Best Local Similarity 99.7%; Pred. No. 6.6e-120;
Matches 308; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 453 TGTGTGGGGGCGAGAGAAACACACACTCACTGCGCCACTTCATCTGTGACCTGTGAG 512
Db 310 TGTGTGGGGGCGAGAGAAACACACACTCACTGCGCCACTTCATCTGTGACCTGTGAG 251

QY 513 GCCCACCCTGCGCTGCTGAGAGAGGCCACAGAGTCCCTTGAATTCGTGACAGCA 572
Db 250 GCCCACCCTGCGCTGCTGAGAGAGGCCACAGAGTCCCTTGAATTCGTGACAGCA 191

QY 573 TGAGATCGTGTGCTGATGAGGGGGCCAGGAGCTGAACCTCTCTGATGACCTGTGAG 632
Db 190 TGAGATCGTGTGCTGATGAGGGGGCCAGGAGCTGAACCTCTCTGATGACCTGTGAG 131

QY 633 CAACATCAACCCGCGACACACCCCAAGGCTGCGGAGAACCTTCACCTTCGTGTGAGAT 692
Db 130 CAACATCAACCCGCGACACACCCCAAGGCTGCGGAGAACCTTCACCTTCGTGTGAGAT 71

QY 693 TTTCACATCATCAAGTCTCTCTCTATCCAGAGCAAGACAGATCATTAATTAATTA 752
Db 70 TTTCACATCATCAAGTCTCTCTCTATCCAGAGCAAGACAGATCATTAATTAATTA 11

QY 753 TGTACTTTA 761
Db 10 TGTACTTTA 2

RESULT 9
BF911380 253 bp mRNA linear EST 18-JAN-2001
LOCUS IL2-UT0073-081100-213-DO2_1 UT0073 Homo sapiens cDNA, mRNA
DEFINITION
ACCESSION BF911380
VERSION BF911380.1 GI:12302838
KEYWORDS
SOURCE EST.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagel, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL
MEDLINE
COMMENT

sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL2&tl2=IL2-UT0073-081100-213-DO2.1&tl3=2000-11-08&tl4=1)
Seq primer: puc18 forward
High quality sequence stop: 220.

FEATURES
source

Location/Qualifiers
1. 253
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lid="pT70073"
/dev_stage="Adult"
/note="Organ: uterus_tumor; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 51 a 67 c 79 g 56 t

ORIGIN

Query Match 30.8%; Score 235; DB 10; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.8e-108;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 GGAACATCAGGGGAGAACATGATCTGATTCCTCTGCTGGGTGGGCGCGAGAAAG 417
Db 1 GGAACATCAGGGGAGAACATGATCTGATTCCTCTGCTGGGTGGGCGCGAGAAAG 60

QY 418 GCTGTGTTCACCTGAGATCTGGAGTGTGAGTGTCTTTGGGGCCAGAGAAACACA 477
Db 61 GCTGTGTTCACCTGAGATCTGGAGTGTGAGTGTCTTTGGGGCCAGAGAAACACA 120

QY 478 CTCACAGCCCACTTATCTGTGACCTGTCTAGAGCCCACTGCGCTGCCCTGAGGA 537
Db 121 CTCACAGCCCACTTATCTGTGACCTGTCTAGAGCCCACTGCGCTGCCCTGAGGA 180

QY 538 GGCCACAGGTCCCTTCTGAAATTCGACACATGAGATGGTGTGATGAG 592
Db 181 GGCCACAGGTCCCTTCTGAAATTCGACACATGAGATGGTGTGATGAG 235

RESULT 10
AM265432 308 bp mRNA linear EST 28-DEC-1999
LOCUS XG52H11.X1 NCI_CGAP_HN7 Homo sapiens cDNA clone IMAGE:2754309 3,
DEFINITION
ACCESSION AM265432
VERSION AM265432.1 GI:6642248
KEYWORDS
SOURCE EST.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?fl=IL3&cl=IL3-UT0114-
301100-356-c03&cl=2000-11-30&cl=1)
Seq primer: puc 18 forward
High quality sequence stop: 245.
Location/Qualifiers
1. 245
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0114"
/dev_stage="Adult"
/note="Organ: uterus; tumor: Vector: puc18; Site: 1: Sma1;
Site: 2: Sma1; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 55 a 63 c 82 g 45 t
ORIGIN

Query Match 20.7%; Score 158; DB 10; Length 245;
Best Local Similarity 99.5%; Pred. No. 3.4e-69;
Matches 208; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 510 GAGGCCACCCCTGCGCTGAGAGAGGCCACAGCTCCCTCTAGATTCTGGACA 569
|||||
DB 209 GAGGCCACCCCTGCGCTGAGAGAGGCCACAGCTCCCTCTAGATTCTGGACA 150
|||||
QY 570 GCATGAGATCGTGTGTGATGGGGGCCAGGAGCTGTGAACCTCTGTATGACCC 629
|||||
DB 149 GCATGAGATCGTGTGTGATGGGGGCCAGGAGCTGTGAACCTCTGTATGACCC 90
|||||
QY 630 GGCACATCATACCCGGGACACCCCAAGCTGGCTGGGAACCTTCACCTCTGTGA 689
|||||
DB 89 GGCACATCATACCCGGGACACCCCAAGCTGGCTGGGAACCTTCACCTCTGTGA 30
|||||
QY 690 GATTTCATCATCTCAAGTCTCTCTTA 718
|||||
DB 29 GATTTCATCATCTCAAGTCTCTCTTA 1
|||||

RESULT 13
BF914311/c 245 bp mRNA linear EST 18-JAN-2001
LOCUS IL3-UT0114-301100-356-E11 UT0114 Homo sapiens CDNA, mRNA sequence.
DEFINITION BF914311
ACCESSION BF914311 GI:12305769
VERSION
KEYWORDS
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 245)
Dias Neto E., Garcia Correia R., Verjovski-Almeida S., Brijones M.R.,
Nagel M.A., da Silva W. Jr., Zagro M.A., Bordin S., Costa F.F.,
Goldman G.H., Carvalho A.F., Matsukuma A., Bala G.S., Simpson D.H.,
Brunstein A., de Oliveira P.S., Bucher P., Jongeneel C.V., O'Hare
M.U., Soares F., Brentani R.R., Reis L.F., de Souza S.J. and
Simpson A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?fl=IL3&cl=IL3-UT0114-
301100-356-E11&cl=2000-11-30&cl=1)
Seq primer: puc 18 forward
High quality sequence stop: 245.
Location/Qualifiers
1. 245
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0114"
/dev_stage="Adult"
/note="Organ: uterus; tumor: Vector: puc18; Site: 1: Sma1;
Site: 2: Sma1; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 55 a 63 c 81 g 46 t
ORIGIN

Query Match 15.2%; Score 116; DB 10; Length 245;
Best Local Similarity 100.0%; Pred. No. 7.2e-48;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 510 GAGGCCACCCCTGCGCTGAGAGAGGCCACAGCTCCCTCTAGATTCTGGACA 569
|||||
DB 209 GAGGCCACCCCTGCGCTGAGAGAGGCCACAGCTCCCTCTAGATTCTGGACA 150
|||||
QY 570 GCATGAGATCGTGTGTGATGGGGGCCAGGAGCTGTGAACCTCTGTATGACCC 625
|||||
DB 149 GCATGAGATCGTGTGTGATGGGGGCCAGGAGCTGTGAACCTCTGTATGACCC 94
|||||

RESULT 14
A1572169 275 bp mRNA linear EST 13-APR-1999
LOCUS A1572169
DEFINITION A1572169
ACCESSION A1572169
VERSION A1572169.1 GI:4535543
KEYWORDS
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 275)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgarbs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 583 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 272.
Location/Qualifiers
1. 275
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:2088898"
/clone_lib="Soares_NhmPc_S1"

/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
 /lab_host="DH10B"
 /note="Organ: mixed (see below); Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2nbhm, pregnant uterus NbHpu, and fetal heart NbH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."
 BASE COUNT 58 a 77 c 56 g 84 t
 ORIGIN
 Query Match 9.7%; Score 74; DB 9; Length 275;
 Best Local Similarity 100.0%; Pred. No. 1.6e-26;
 Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 121 ACCGTATGCTGTCAGCCACACAGAGATGTGGGACAGTTCTACGACCCCTGCAC 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 178 ACCGTATGCTGTCAGCCACACAGAGATGTGGGACAGTTCTACGACCCCTGCAC 237
 QY 181 ACTGTGCTATGAT 194
 ||||||||||||||||
 Db 238 ACTGTGCTATGAT 251

RESULT 15

BG951874/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 25 a 39 c 22 g 17 t
 ORIGIN
 Query Match 7.9%; Score 60; DB 10; Length 103;
 Best Local Similarity 100.0%; Pred. No. 1.6e-19;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 366 AGGGAGACGATGACTGCTGATTCCTCTCTGGTGGGCGTGGAGAAAGAGCGTGT 425
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 103 AGGGAGACGATGACTGCTGATTCCTCTCTGGTGGGCGTGGAGAAAGAGCGTGT 44
 Search completed: November 3, 2002, 03:02:14
 Job time : 1698 secs

FEATURES

Source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT0663"
 /dev_stage="Adult"

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1161
US-09-389-528-22

Query Match
Best Local Similarity 100.0%; Score 17; DB 4; Length 1161;
Pred. No. 35;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CCGTGGTGGCCCTGGCC 215
DB 350 CCGTGGTGGCCCTGGCC 366

RESULT 14
US-09-181-827A-22
Sequence 22, Application US/09181827A
Patent No. 6242232
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie S.
APPLICANT: Brown, Kimberly M.
APPLICANT: Kauppinen, Sakari
APPLICANT: Halkier, Torben P.
TITLE OF INVENTION: Purified Coprinus Laccases And Nucleic
FILE REFERENCE: 4554,200-US
CURRENT APPLICATION NUMBER: US/09/181,827A
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/002,800
PRIOR FILING DATE: 1995-08-25
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22
LENGTH: 1161
TYPE: DNA
ORGANISM: Coprinus cinereus
US-09-181-827A-22

Query Match
Best Local Similarity 100.0%; Score 17; DB 4; Length 1161;
Pred. No. 35;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 CCGTGGTGGCCCTGGCC 215
DB 350 CCGTGGTGGCCCTGGCC 366

RESULT 15
US-08-485-284A-3
Sequence 3, Application US/08485284A
Patent No. 5750372
GENERAL INFORMATION:
APPLICANT: SAKAI, YASUYOSHI
APPLICANT: TANI, YOSHIKI
APPLICANT: SHIBANO, YUJI
APPLICANT: KONDO, HIROMO
APPLICANT: HATANAKA, HARUYO
TITLE OF INVENTION: NOVEL VECTOR HAVING PROMOTER THAT IS
INDUCIBLE BY METHANOL AND/OR GLYCEROL
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D. C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,284A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 43361/1992
FILING DATE: 28-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/025,416
FILING DATE: 01-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: WHITE JR., PAUL E.
REGISTRATION NUMBER: 32,011
REFERENCE/DOCKET NUMBER: 217755/FPS382090S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 1992 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: genomic DNA
US-08-485-284A-3

Query Match
Best Local Similarity 100.0%; Score 17; DB 1; Length 1992;
Pred. No. 35;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 283 TGGTGAAGCTGATTAAC 299
DB 1563 TGGTGAAGCTGATTAAC 1579

Search completed: November 3, 2002, 01:07:56
Job time : 59 secs

LOCATION: (1)...(570)
US-08-884-866A-2

Query Match 2.2%; Score 17; DB 4; Length 570;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 606 CTGACCCCTCGATGA 622
|||||

DB 511 CTGACCCCTCGATGA 527

RESULT 11

US-08-884-866A-11
Sequence 11, Application US/08884866A

GENERAL INFORMATION:

APPLICANT: Chien, John Y-J

TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND

TITLE OF INVENTION: BYPASS

FILE REFERENCE: UCSD1100-1

CURRENT APPLICATION NUMBER: US/08/884,866A

CURRENT FILING DATE: 1997-06-30

PRIOR APPLICATION NUMBER: 60/030,358

PRIOR FILING DATE: 1996-11-08

NUMBER OF SEQ ID NOS: 25

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11

LENGTH: 570

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: misc-feature

LOCATION: (49)...(51)

OTHER INFORMATION: nm - Any nucleic acid triplet, except for UCA,

OTHER INFORMATION: UCC, UCG, UCU, AGC, or AGU

OTHER INFORMATION: Variation of SEQ ID NO.: 2

Query Match 2.2%; Score 17; DB 4; Length 570;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 606 CTGACCCCTCGATGA 622
|||||

DB 511 CTGACCCCTCGATGA 527

RESULT 12

US-08-689-421-22

Sequence 22, Application US/08689421

Patent No. 6008029

GENERAL INFORMATION:

APPLICANT: Yaver, Debbie S.

APPLICANT: Brown, Kimberley M.

APPLICANT: Kaupinen, Sakari

APPLICANT: Halkier, Torben P

TITLE OF INVENTION: PURIFIED COPRINUS LACCASES AND NUCLEIC ACIDS

TITLE OF INVENTION: ENCODING SAME

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESS: No. 60080290 No. 6008029disk of No. 6008029th America, Inc.

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/689,421

FILING DATE: 9-AUG-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Harrington, James J.

REGISTRATION NUMBER: 38,711

REFERENCE/DOCKET NUMBER: 4554, 204-WO

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 1161 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1161

Query Match 2.2%; Score 17; DB 3; Length 1161;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 199 CCGTGGCCCTGGCC 215
|||||

DB 350 CCGTGGCCCTGGCC 366

RESULT 13

US-09-389-528-22

Sequence 22, Application US/09389528

Patent No. 6207430

GENERAL INFORMATION:

APPLICANT: Yaver, Debbie S.

APPLICANT: Brown, Kimberley M.

APPLICANT: Kaupinen, Sakari

APPLICANT: Halkier, Torben P

TITLE OF INVENTION: PURIFIED COPRINUS LACCASES AND NUCLEIC ACIDS

TITLE OF INVENTION: ENCODING SAME

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESS: No. 62074300 No. 6207430disk of No. 6207430th America, Inc.

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/389,528

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/689,421

FILING DATE: 9-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: Harrington, James J.

REGISTRATION NUMBER: 38,711

REFERENCE/DOCKET NUMBER: 4554, 204-WO

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 1161 base pairs

TYPE: nucleic acid

;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-08-256-368-5

Query Match 2.2%; Score 17; DB 1; Length 45;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 551 CCTCTAGAAATCTGGA 567
|||||
DB 38 CCTCTAGAAATCTGGA 22

RESULT 8

US-08-256-368-6
; Sequence 6, Application US/08256368
; Patent No. 5641649
; GENERAL INFORMATION:
; APPLICANT: Stanchl, Ombretta
; APPLICANT: Negro, Alessandro
; APPLICANT: Callegaro, Lanfranco
; TITLE OF INVENTION: Expression Of Osteogenic Factor Op-1
; TITLE OF INVENTION: In Cells Of Spodoptera Frugiperda Infected With
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,368
; FILING DATE: 22-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 259-243P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)205-8000
; TELEFAX: (703)205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = complementary 3'-5' strand of synthetic
; DESCRIPTION: polylinker in pVL1392. See Figure 2."
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-256-368-6

Query Match 2.2%; Score 17; DB 1; Length 45;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 551 CCTCTAGAAATCTGGA 567
|||||
DB 12 CCTCTAGAAATCTGGA 28

RESULT 9
US-08-716-308-9/c

;; Sequence 9, Application US/08716308
;; Patent No. 5885369
;; GENERAL INFORMATION:
;; APPLICANT: Windass, John D.
;; TITLE OF INVENTION: Biological Insect Control Agent
;; NUMBER OF SEQUENCES: 18
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ZENECA INC.
;; STREET: 1800 Concord Pike
;; CITY: Wilmington
;; STATE: DE
;; COUNTRY: USA
;; ZIP: 19850

COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/716,308
;; FILING DATE: 24-SEP-1996
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/GB95/00677
;; FILING DATE: 27-MAR-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9405951.6
;; FILING DATE: 25-MAR-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hohenschutz, Liza D.
;; REGISTRATION NUMBER: 33,712
;; REFERENCE/DOCKET NUMBER: PPD40027X/UST
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (302) 886-1699
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 45 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA
US-08-716-308-9

Query Match 2.2%; Score 17; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 551 CCTCTAGAAATCTGGA 567
|||||
DB 38 CCTCTAGAAATCTGGA 22

RESULT 10

US-08-884-866A-2
; Sequence 2, Application US/08884866A
; GENERAL INFORMATION:
; APPLICANT: Chien, Shu
; APPLICANT: Shyy, John Y.-J.
; TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
; TITLE OF INVENTION: BYPASS
; FILE REFERENCE: UCSD1100-1
; CURRENT APPLICATION NUMBER: US/08/884,866A
; CURRENT FILING DATE: 1997-06-30
; PRIOR APPLICATION NUMBER: 60/030,358
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS

APPLICANT: Rothe, Mike
APPLICANT: Wu, Lin
TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,321
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/887,518
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: T97-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3156 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-023-321-1
Query Match 2.4%; Score 18; DB 2; Length 3156;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 225 ACGTGTGGAAGTGCACC 242
DB 713 ACGTGTGGAAGTGCACC 730
RESULT 6
US-09-032-475-1
Sequence 1, Application US/09032475
Patent No. 5854003
GENERAL INFORMATION:
APPLICANT: Rothe, Mike
APPLICANT: Wu, Lin
TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032,475
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/887,518
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: T97-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3156 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-032-475-1
Query Match 2.4%; Score 18; DB 2; Length 3156;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 225 ACGTGTGGAAGTGCACC 242
DB 713 ACGTGTGGAAGTGCACC 730
RESULT 7
US-08-256-368-5/c
Sequence 5, Application US/08256368
Patent No. 5641649
GENERAL INFORMATION:
APPLICANT: Starch, Ombretta
APPLICANT: Negro, Alessandro
TITLE OF INVENTION: Expression Of Osteogenic Factor Op-1
TITLE OF INVENTION: In Cells Of Spodoptera Frugiperda Infected With
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256,368
FILING DATE: 22-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Svensson, Leonard R
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 259-243P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)205-8000
TELEFAX: (703)205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "5'-3' strand of synthetic
polylinker in pVL1392. See Figure 2."

Db 387 ACCTATGACATCAGG 405

RESULT 2
US-08-470-179-59

; Sequence 59, Application US/08470179
; Patent No. 5645994

GENERAL INFORMATION:

; APPLICANT: Huang Ph.D, Wai Mun
; TITLE OF INVENTION: Method and Compositions for
; TITLE OF INVENTION: Identification of Species in a Sample
; NUMBER OF SEQUENCES: 207

CORRESPONDENCE ADDRESS:

; ADDRESSEE: Trask, Britt and Rossa
; STREET: P.O. Box 2550
; CITY: Salt Lake City
; STATE: Utah
; COUNTRY: USA
; ZIP: 84110

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,179
; FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

; NAME: Sweigert Ph.D, Susan E.
; REGISTRATION NUMBER: 36,289
; REFERENCE/DOCKET NUMBER: 2601
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 801-532-1922

INFORMATION FOR SEQ ID NO: 59:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "ytrA gene segment"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: Chlamydia trachomatis
; US-08-470-179-59

Query Match 2.4%; Score 18; DB 1; Length 422;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 719 TCCAGAGCAAGCAGG 736
Db 60 TCCAGAGCAAGCAGG 77.

RESULT 3

US-09-257-703-2

; Sequence 2, Application US/09257703

; Patent No. 6265338

GENERAL INFORMATION:

; APPLICANT: Greene, Warner C.

; APPLICANT: Lin, Xin

; APPLICANT: Gelezunas, Romas

; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED

; FILE REFERENCE: 30448-61US01

; CURRENT APPLICATION NUMBER: US/09/257,703

; CURRENT FILING DATE: 1999-02-25

; EARLIER APPLICATION NUMBER: 60/076,299

; EARLIER FILING DATE: 1998-02-27

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 2

; LENGTH: 2844

; TYPE: DNA

; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)

US-09-257-703-2

Query Match 2.4%; Score 18; DB 4; Length 2844;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 ACGTGTGGAAGTGCACC 242
Db 713 ACGTGTGGAAGTGCACC 730

RESULT 4

US-08-887-518-1

; Sequence 1, Application US/08887518

; Patent No. 5843721

GENERAL INFORMATION:

; APPLICANT: Rotne, Mike

; APPLICANT: Wu, Lin

; TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/887,518
; FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: T97-008

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 343-4342

; TELEFAX: (415) 343-4341

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3156 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-08-887-518-1

Query Match 2.4%; Score 18; DB 2; Length 3156;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 ACGTGTGGAAGTGCACC 242
Db 713 ACGTGTGGAAGTGCACC 730

RESULT 5

US-09-023-321-1

; Sequence 1, Application US/09023321

; Patent No. 5844073

GENERAL INFORMATION:

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 23:18:40 ; Search time 48 Seconds
(without alignments)
3904.549 Million cell updates/sec

Title: US-09-092-297-5

Perfect score: 763
Sequence: 1 CTCGACTGCAACCCACCA.....ATTAATTATGACTTATA 763

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 38353 seqs, 12816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/lna/5A_COMB.seq:*
2: /cgn2_6/prodata/1/lna/5B_COMB.seq:*
3: /cgn2_6/prodata/1/lna/6A_COMB.seq:*
4: /cgn2_6/prodata/1/lna/6B_COMB.seq:*
5: /cgn2_6/prodata/1/lna/PCUS_COMB.seq:*
6: /cgn2_6/prodata/1/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	2.5	423	1	US-08-470-179-142 Sequence 142, Appl
2	18	2.4	422	1	US-08-470-179-59 Sequence 59, Appl
3	18	2.4	2844	4	US-09-257-703-2 Sequence 2, Appl
4	18	2.4	3156	2	US-08-887-518-1 Sequence 1, Appl
5	18	2.4	3156	2	US-09-023-321-1 Sequence 1, Appl
6	18	2.4	3156	2	US-09-032-475-1 Sequence 1, Appl
7	17	2.2	45	1	US-08-256-368-5 Sequence 5, Appl
8	17	2.2	45	1	US-08-256-368-6 Sequence 6, Appl
9	17	2.2	45	2	US-08-716-308-9 Sequence 9, Appl
10	17	2.2	570	4	US-08-884-866A-2 Sequence 2, Appl
11	17	2.2	570	4	US-08-884-866A-11 Sequence 11, Appl
12	17	2.2	1161	3	US-08-689-421-22 Sequence 22, Appl
13	17	2.2	1161	3	US-09-389-528-22 Sequence 22, Appl
14	17	2.2	1161	4	US-09-181-827A-22 Sequence 22, Appl
15	17	2.2	1992	1	US-08-485-284A-3 Sequence 3, Appl
16	17	2.2	2238	2	US-08-919-624-2 Sequence 2, Appl
17	17	2.2	2238	2	US-08-919-624-2 Sequence 3, Appl
18	17	2.2	2940	3	US-09-389-528-28 Sequence 28, Appl
19	17	2.2	2940	4	US-09-389-528-28 Sequence 28, Appl
20	17	2.2	3715	4	US-09-181-827A-28 Sequence 28, Appl
21	17	2.2	3715	4	US-09-041-886-10 Sequence 10, Appl
22	17	2.2	3728	1	US-08-111-939-1 Sequence 1, Appl
23	17	2.2	6453	3	US-08-306-691B-14 Sequence 14, Appl
24	17	2.2	6453	3	US-09-209-668-10 Sequence 10, Appl
25	17	2.2	6453	3	US-09-356-952-8 Sequence 8, Appl
26	16	2.1	96	4	US-08-978-741-5 Sequence 5, Appl
27	16	2.1	96	4	US-09-390-867A-19 Sequence 19, Appl
					US-09-548-260-19 Sequence 19, Appl

28	16	2.1	536	4	US-08-991-189A-252	Sequence 252, App
29	16	2.1	536	4	US-09-062-451-252	Sequence 252, App
30	16	2.1	613	4	US-09-385-982-144	Sequence 144, App
31	16	2.1	944	4	US-09-386-493-4	Sequence 4, Appl
32	16	2.1	1152	3	US-09-188-930-42	Sequence 42, Appl
33	16	2.1	1152	3	US-09-188-930-211	Sequence 211, App
34	16	2.1	1421	3	US-09-188-930-70	Sequence 70, Appl
35	16	2.1	1421	3	US-09-188-930-254	Sequence 6, Appl
36	16	2.1	1421	3	US-08-548-159-6	Sequence 1, Appl
37	16	2.1	3515	2	US-08-391-743A-1	Sequence 1, Appl
38	16	2.1	4035	6	5198359-1	Patent No. 5198359
39	16	2.1	4035	6	5449756-1	Patent No. 5449756
40	16	2.1	4182	1	US-07-973-257-1	Sequence 1, Appl
41	16	2.1	7360	4	US-09-103-478-4	Sequence 4, Appl
42	16	2.1	7560	4	US-09-193-931C-4	Sequence 254, Appl
43	16	2.1	48974	4	US-08-920-422-17	Sequence 17, Appl
44	16	2.1	246240	2	US-08-724-394A-20	Sequence 20, Appl
45	16	2.1	246240	2	US-08-724-394A-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-470-179-142
Sequence 142, Application US/08470179
Patent No. 5645994
GENERAL INFORMATION:
APPLICANT: Huang Ph.D, Wei Mun
TITLE OF INVENTION: Method and Compositions for
IDENTIFICATION OF SPECIES IN A SAMPLE
NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Task, Britt and Rossa
STREET: P.O. Box 2550
CITY: Salt Lake City
STATE: Utah
COUNTRY: USA
ZIP: 84110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,179
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sweigert Ph.D, Susan E.
REGISTRATION NUMBER: 36,289
REFERENCE/DOCKET NUMBER: 2601
TELECOMMUNICATION INFORMATION:
TELEPHONE: 801-532-1922
TELEFAX: 801-531-9168
INFORMATION FOR SEQ ID NO: 142:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mycoplasma arthritidis
US-08-470-179-142
Query Match 2.5%: Score 19; DB 1; Length 423;
Best Local Similarity 100.0%; Pred No. 3.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 351 AGCTATGACATCAGG 369

```

/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42043.1"
/db_xref="GI:7226950"
/translation="MELQORTLTIGLNDIAVQTEGVGFIKRGOLVADHAENSYFKV
VKALIVKTEFADGYPHAIFORFLMDNRNRKRAKAGVLGADDAEIRICLI
SNAPFTELLIGICGKADCGHEFVVAASASVAKVSDQNEFAVRFPAGKOHPOY
IGNHAGINRPPIISPOCQSGKRPVDYKRYAPNLRYPAGIAGFFRQVCLQPKRF
NRLPGQAAAEHROHDIQEDGIPRIHMPPEHGYCKKQCNKNSRITLLGQSGG
AOCPOPARHRQPKPARLPAYVSGLTSTALPRLAVLSASSPCBDFC"
complement(4988..5233)
/gene="NMB1696"
complement(4988..5233)
/note="similar to PID:870808 percent identity: 58.23;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="acyl carrier protein"
/protein_id="AAF42044.1"
/db_xref="GI:7226951"
/translation="MTEQETRLRLDTLTLEFIEPERIPDNLVDELDISDAID
LIDRIKREGRKIQAEDEFNRVNDVQAVLKIQAG"
complement(5245..5505)
/gene="NMB1697"
complement(5245..5505)
/note="similar to PID:1655811 percent identity: 52.94;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="acyl carrier protein, putative"
/protein_id="AAF42045.1"
/db_xref="GI:7226952"
/translation="MMDLENOIKQLIIDSALEDITRADIGSFDAIFGQGLGLDSVD
ALEGLAVQKHREFRLDGEQENLREHFANKVTLAAVKSQA"
complement(5502..6260)
/gene="NMB1698"
complement(5502..6260)
/note="similar to GB:AE000511 PID:2314519 percent
identity: 52.85; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="acyltransferase, putative"
/protein_id="AAF42046.1"
/db_xref="GI:7226953"
/translation="MDKLDYCRFFATWLFVIFGVGIMKLVLLPYTLNLTSGSVA
ROLAARRIIGTSWRLFAVILKNSGVLEFVKGEKLNREGOLILAHSPGLLVYLVG
HYPEMNCIVKDLQHNPAISOIKGAGYIPNSESSEMLTTRAVPDSQSLVPEPGT
RTGMDGRVYKMGAVSIGLRVAEVLTPVCIKMNPFKKGQWYRIRPRIRIETIVG
DILPODWLAEKPLPAARLNELVLDYTRKTI"
complement(6253..6618)
/gene="NMB1699"
complement(6253..6618)
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42047.1"
/db_xref="GI:7226954"
/translation="MSEQFALAVCADGVETALAAEASLIEGCGSVLYIADDPLEPG
YAVSATRAPFAALAMVLTGRTISLTLSASDMPSEAGMLPEAYWGLGVVRFLLNG
SRECRRVIRNRKEMWQRNG"
complement(6778..7143)
/gene="NMB1700"
complement(6778..7143)

```

```

/gene="NMB1700"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42048.1"
/db_xref="GI:7226955"
/translation="MGGANVPSRVAQOTRLTOTAAHSGOERQRPDQGVKIRF
FRPCPLGHIHAFARNDPGGDEVEEANNRRHSFPQKFTLCRLITKAYPPDNAEIQ
HFCPLSTTPELCKPPLA"
complement(7196..7696)
/gene="NMB1701"
complement(7196..7696)
/note="NMB1701"
/gene="NMB1701"
complement(7196..7696)
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42049.1"
/db_xref="GI:7226956"
/translation="MDPLPCPTNTASLPHSGRMVLIBRITRYGDFVEAGQVSP
NHILLDDKLPYTAFTILMAQVAYGIAQRNARSVRIGLGTAKLIFASQVPI
GTHLATNAHSIODAGMGVDELRMTDAPETSSETLPSDGIILARASLNVSPENHA
GTHDAV"
complement(7724..8452)
/gene="NMB1702"
complement(7724..8452)
/note="NMB1702"
/note="similar to SP:P70720 PID:1651214 percent identity:
90.91; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="3-oxoacyl-(acyl-carrier-protein) reductase"

```

Alignment Scores:

Pred. No.:	Score:	Length:
210	8.00	12050
Percent Similarity:	100.00%	Matches:
Best Local Similarity:	100.00%	Conservative:
Query Match:	6.84%	Mismatches:
DB:	1	Indels:
		Gaps:

US-09-092-297-17 (1-117) x AE002520 (1-12050)

QY 45 GlnProHisLysArgCysGlyASP 52

DB 8356 CAGCCCATTAAGCGCTGCGGGGAC 8333

Search completed: November 3, 2002, 05:46:48
Job time: 2012 secs

US-09-092-297-17 (1-117) x HSCALCAC (1-7637)

QY 90 PROTRPThPheMeValylsLeu 97
 DB 6265 CCATGACCTTCATGTAAGTTG 6242

RESULT 14
 AC017452
 LOCUS
 DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***; in ordered pieces.
 AC017452
 AC017452.1 GI:5553534
 HMG: HMG5.PHASE2.
 fruit fly.
 ORGANISM Drosophila melanogaster.
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
 1 (bases 1 to 9366)
 Adams,M. and Venter,J.C.
 TITLE Direct Submission
 JOURNAL Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

COMMENT This sequence was identified as CDM:10210836 by the submitter. For more information on this record e-mail to fly@celera.com.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
 source
 1..9366
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"

BASE COUNT 2736 a 2089 c 2024 g 2517 t

ORIGIN
 Alignment Scores:
 Pred. No.: 165 Length: 9366
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.84% Indels: 0
 DB: 2 Gaps: 0

US-09-092-297-17 (1-117) x AC017452 (1-9366)

QY 21 PheCysIleSerArgLeuLeuCys 28
 DB 6399 TTCGCATTTCTCGATGTATGTC 6422

RESULT 15
 AE002520/c
 LOCUS
 DEFINITION Neisseria meningitidis serogroup B strain MC58 section 162 of 206 of the complete genome.
 AC002520 AE002520
 VERSION AE002520.1 GI:7226947
 KEYWORDS
 SOURCE
 ORGANISM
 Neisseria meningitidis MC58.
 Neisseria meningitidis MC58
 Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;

REFERENCE
 1 (bases 1 to 12050)
 Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., Deboy,R., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E., Citrono,H., Clark,E.B., Cotton,M.D., Uitterback,T.R., Khouri,H., Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Maignani,V., Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,

TITLE Rappuoli,R. and Venter,J.C.
 JOURNAL Complete genome sequence of Neisseria meningitidis serogroup B strain MC58
 MEDLINE Science 287 (5459), 1809-1815 (2000)
 PUBMED 20175755
 10710307

REFERENCE
 2 (bases 1 to 12050)
 Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., Deboy,R., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E., Citrono,H., Clark,E.B., Cotton,M.D., Uitterback,T.R., Khouri,H., Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Maignani,V., Rappuoli,R. and Venter,J.C.
 TITLE Direct Submission
 JOURNAL Submitted (17-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

FEATURES
 source
 1..12050
 /organism="Neisseria meningitidis MC58"
 /strain="MC58"
 /db_xref="taxon:122586"
 /note="serogroup: B"
 76..2190
 /gene="NMB1693"
 76..2190
 /gene="NMB1693"
 /note="NMB1693"
 /note="hypothetical protein; identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=1
 /product="hypothetical protein"
 /protein_id="AA042041.1"
 /db_xref="GI:7226948"
 /translation="MDLSVFKRYRLKAVAVLITLLAVGASHVRYTPFPENTRS
 RLODSIAHTRKISFDADIORLLRPPTVLKMLTTEPGDQTVASVOETKIGLSMK
 LMMSDOIOEKWVSSAEALATRDGQVNWIOIDSKQASVNIIVYNSRNF
 LOEOLIKETINLMSDSSGQPPSSGILVCKLISPKRSRLTSLNKGPPSEIPE
 HFPASTSLDGHGTTISTGSPSYRPNAGDAAGLRADTSFRNLHQTPLALR
 NNSKIKETTVNGATGGEIVARWDSFELDKANLHSGIANGNEISGSEKTPRHQTF
 SLNSPLVTENKGLDARLVSTLQDVTNLPQPRISRDGSLSVPNQNMVAELNG
 TFDQRTVAAKFRYTHEDAPHLAVALQKMLTPYLDVYQCGKGFEPITLALSGDI
 EAHKIKGVQLPGLQDDMETYLHADGHALSRFSGLYGGTEGGISIANRPATY
 RLOONASNIQIOLPDLDFEGFHSFGSGDAVIDLTAGGGRKELISIGSGSINTSN
 GAWHGIDMDNIITKNGISGRADNAASPTPHRTLSNSEDGSIIRITDIBSDSLTY
 TSNGYTRLDIOELSEDEVLLIRNAVHPKRPILPKITGTVDPSITVDYGRLTGINSRK
 EKQRILEDYDLEQWMLKPKRP"
 /complement(2254..3738)
 /gene="NMB1694"
 /complement(2254..3738)
 /gene="NMB1694"
 /note="conserved hypothetical protein; identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=1
 /product="conserved hypothetical protein"
 /protein_id="AA042042.1"
 /db_xref="GI:7226949"
 /translation="NNMKIKIDLNFLEAMEQGLKRVAPHSPLYEMTEIADRYLRA
 EGPALLFENPIKPDGTGRGYVLANLEGPERRVAMGMSVSKLEIQTLAYLKEP
 EPPKGIKDAFSKPLKLDLMSNAPVNVKNAPOCEIYWESEGDVLYLPLOHCPEDVA
 PLVTGCLVYTRGPHKRONIGYRQOLIGKNTKIMVLSHSGALDYOGRRLNDPT
 YPAAVILGDDPATIIGAVTPVPTLSEYQFAGILRSRPELTKICINDQVAPRAEIV
 LCVTHIPNTEALEGPGDHTGYNEDDYTPVTVEITRNPITHSTYGRPPDEPA
 VGVANAEVFPVLQGFPEIDETLPEGSCSTRAAVSMKQVAGHARVAGGWSF
 LRGFTMTKTIIVDDDVNDRMKVIMAVTTRDDPVKDVLENTPIDYIDLASPVSG
 LGQKMLDANKPKPGETDRENGRVIKKDAVYAKIDGIMEELG"
 3857..4837
 /gene="NMB1695"
 3857..4837
 /gene="NMB1695"
 /gene="NMB1695"

```

FEATURES
SOURCE
Location/Qualifiers
1. 2350
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
join(<501. .1199,1557. .>1883)
/product="putative metal transporter ZIP6"
join(501. 1199,1557. .1883)
/codon_start=1
/product="putative metal transporter ZIP6"
/protein_id="AAL38433.1"
/db_xref="GI:17385786"
/translation="MASCVMGTEAATRAACROGSEASHUKIYAVEAIFLTSVGVGAC
PVLLAKFHGKPLVDKALIVIKCFRAAYILSTSLVHYLPFAFESLADCOVSRHPKWD
PFRGLVYTMIGATILALVDLPASEHGGGGGDMETPVKAVAGLEMEGKGG
ADLEIENSEEELIYKMLRVDSOYLEIGIIFHSYIIEVTGMSQNKCTIRPLIALSF
HOIEPGLGIGGICIAOGCFKAGTVYVIMLNAVTPPLIVGMVFEATGYDDQNPML
IMEILGFSFGSILIIYVALVDLIAIDFHNKMLITCGESSRLKILCFVALYGSAMN
SLALVLA"
BASE COUNT 707 a 425 c 475 g 743 t
ORIGIN
Alignment Scores:
Pred. No.: 44.1 Length: 2350
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.84% Indels: 0
DB: 8 Gaps: 0
US-09-092-297-17 (1-117) x AF369910 (1-2350)
Oy 14 11levala1a1phea1a1lephe 21
Db 582 ATAGTCGCCGCTTCGCCATCTTT 605
RESULT 13
LOCUS HSCALCAC 7637 bp DNA linear PRI 25-JUN-1997
DEFINITION Human calcitonin/alpha-CGRP gene.
ACCESSION X15943
VERSION X15943.1 GI:29613
KEYWORDS alternate splicing; Alu repetitive sequence; calcitonin; calcitonin
gene related peptide; hormone; neuropeptide; repetitive sequence.
human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 7637)
AUTHORS Broad,P.M., Symes,A.J., Thakker,R.V. and Craig,R.K.
TITLE Structure and methylation of the human calcitonin/alpha-CGRP gene
JOURNAL Nucleic Acids Res. 17 (17), 6999-7011 (1989)
MEDLINE 89386053
REFERENCE 2 (bases 1 to 7637)
AUTHORS Symes,A.J.
TITLE Direct Submission
JOURNAL Submitted (22-JUL-1989) Symes A.J., University College London,
Medical Molecular Biology Unit, Windeyer Building, UCMWS, Cleveland
street, London W1P 6DB, UK
FEATURES
SOURCE
Location/Qualifiers
1. 6937
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="P15.4"
/clone_id="cosmid_HPB-ALL cos 202"
1573. .1582
/note="c-AMP responsive element"
1666. .1674
/note="octamer element"
1752. .1757
/note="spi binding site"
/bound_moiety="Spi"
1804. .1808
TATA_signal

```

MRNA	join(1834, .1940,2983, .3077,4108, .4248,5188, .5682, 7085, .7515)	/gene="CGRP"
	/product="calcitonin"	
prim_transcript	1834, .7515	
	/gene="CGRP"	
gene	/note="calcitonin/CGRP gene transcript"	
	join(1834, .1940,2983, .3077,4108, .4248,5188, .5682, 7085, .7515)	
exon	/gene="CGRP"	
	1834, .1940	
	/gene="CGRP"	
intron	/note="exon 1, common"	
	1942, .2982	
	/gene="CGRP"	
exon	/note="intron I"	
	2983, .3077	
	/gene="CGRP"	
CDS	/note="exon 2, common"	
	join(2992, .3077,4108, .4248,6330, .6489)	
	/gene="CGRP"	
	/codon_start=1	
	/product="cgip propeptide"	
	/protein_id="CAA34070.1"	
	/db_xref="GI:296638"	
	/db_xref="SWISS-PROT:P01258"	
	/db_xref="EMBL:P06881"	
	/translation="MGFOKSFPLALSLVLLGSLHAAPFRSALESPADPATLSE DEALLAALVODVOMKASELEQBRBSRIIAQKRAQDTATCVTHRLAGLSRSG GVYNNFVPTNVGSKAFRRRRDQA"	
sig_peptide	2992, .3066	
	/gene="CGRP"	
mat_peptide	join(3067, .3077,4108, .4248,5188, .5311)	
	/gene="cgip"	
gene	/product="calcitonin"	
	join(3067, .3077,4108, .4248,5188, .5311)	
intron	/gene="cgip"	
	3078, .4107	
	/gene="cgip"	
	/note="intron II"	
repeat_region	3287, .3568	
	/note="Alu repeat"	
exon	4108, .4248	
	/gene="cgip"	
intron	/note="exon 3, common"	
	4249, .5187	
	/gene="cgip"	
exon	/note="intron III"	
	5188, .5682	
	/gene="CGRP"	
intron	/note="exon 4, specific to preprocalcitonin"	
	5683, .6329	
	/gene="CGRP"	
exon	/note="intron IV"	
	6330, .6510	
	/gene="CGRP"	
intron	/note="exon 5, specific to CGRP"	
	6511, .7084	
	/gene="CGRP"	
exon	/note="intron 5"	
	7085, .7515	
	/gene="CGRP"	
	/note="exon 6"	
BASE COUNT	1843 a 1939 c 1922 g 1933 t	
ORIGIN		
Alignment Scores:		
Pred. No.:	136	Length: 7637
Score:	8.00	Matches: 8
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	6.84%	Indels: 0
DB:	9	Gaps: 0

Tag Poly: 0.05 units/ul
Total Vol: 15 ul

Buffer: 1.5mM
MgCl2: 50 mM
KCl: 50 mM
Tris-HCl: 8.3
pH: .001 %
Gelatin:

P-element stocks should be obtained from the IV Drosophila Stock Center.

FEATURES

source location/Qualifiers
1..329
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
STS
primer_bind 49..201
primer_bind 49..68
complement(180..201)
BASE COUNT 109 a 54 c 59 g 106 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 6.75 Length: 329
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.84% Indels: 0
DB: 11 Gaps: 0

US-09-092-297-17 (1-117) x G00569 (1-329)

OY 29 SerHsgAlaProValAlaPro 36
Db 64 TCTCATGCTGCACCTGTGACACGC 41

RESULT 10

AX309824/c AX309824 373 bp DNA linear PAT 14-DEC-2001
LOCUS Sequence 2809 from Patent WO0190366.
DEFINITION AX309824
ACCESSION AX309824
VERSION AX309824.1 GI:17896142
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS Leach,M.D. and Shinkets,R.A.
TITLE Human polynucleotides and polypeptides encoded thereby
JOURNAL Patent: WO 0190366-A 2809 29-NOV-2001;
Curagen Corporation (US)
location/Qualifiers

FEATURES
source 1..373
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 94 a 105 c 97 g 76 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 7.61 Length: 373
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.84% Indels: 0
DB: 6 Gaps: 0

US-09-092-297-17 (1-117) x AX309824 (1-373)

OY 103 AspSerAlaArgThrSerAspASP 110
Db 41 GACTCTGCCGCTTCTGTATGAC 18

RESULT 11

AF320338 792 bp DNA linear BCT 24-JAN-2001
LOCUS AF320338/c
DEFINITION uncultured alpha proteobacterium rcr_1 16S ribosomal RNA gene,
partial sequence.
ACCESSION AF320338
VERSION AF320338.1 GI:12407987
KEYWORDS
SOURCE uncultured alpha proteobacterium rcr_1.
ORGANISM uncultured alpha proteobacterium; alpha subdivision; environmental samples.
REFERENCE 1 (bases 1 to 792)
AUTHORS Boomer,S.M., Dutton,B.E. and Lodge,D.P.
TITLE Diversity of novel green non-sulfur bacteria from Yellowstone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 792)
AUTHORS Boomer,S.M., Dutton,B.E. and Lodge,D.P.
TITLE Direct Submission
JOURNAL Submitted (09-NOV-2000) Dept. of Biology, Western Oregon
University, 345 Monmouth Ave., Monmouth, OR 97361, USA
location/Qualifiers

FEATURES

source 1..792
/organism="uncultured alpha proteobacterium rcr_1"
/db_xref="taxon:148390"
/clone="rcr_1"
/country="USA: Rabbit Creek, Yellowstone"
complement(<1..>792)
product="16S ribosomal RNA"
BASE COUNT 166 a 233 c 204 g 189 t
ORIGIN

Alignment Scores:

Pred. No.: 15.6 Length: 792
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.84% Indels: 0
DB: 1 Gaps: 0

US-09-092-297-17 (1-117) x AF320338 (1-792)

OY 1 ProLeuGlnProProArgAlaMet 8
Db 282 CCTCTGCAACCTCGAGGCCATG 259

RESULT 12

AF369910 2350 bp DNA linear PLN 06-DEC-2001
LOCUS AF369910
DEFINITION Arabidopsis thaliana putative metal transporter ZIP6 gene, complete
cds.
ACCESSION AF369910
VERSION AF369910.1 GI:17385785
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 2350)
Maser,P., Thomine,S., Schroeder,J.I., Ward,J.M., Hirschl,K.,
Sez,H., Talke,I.N., Amtmann,A., Maathuis,F.J., Sanders,D.,
Harper,J.F., Tchileu,J., Gribkov,M., Persans,M.W., Salt,D.E.,
Kim,S.A. and Gueriot,M.L.
TITLE Phylogenetic relationships within cation transporter families of
Arabidopsis
JOURNAL Plant Physiol. 126 (4), 1646-1667 (2001)
MEDLINE 21392307
PUBMED 11500563
REFERENCE 2 (bases 1 to 2350)
AUTHORS Kim,S.A. and Gueriot,M.L.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2001) Department of Biology, Dartmouth College,
Hanover, NH 03755, USA

```

* 124037 124136: gap of unknown length
* 124137 134245: contig of 10109 bp in length
* 134246 134345: gap of unknown length
* 134346 143789: contig of 9444 bp in length
* 143790 143889: gap of unknown length
* 143890 153881: contig of 11492 bp in length
* 153882 155481: gap of unknown length
* 155482 159655: contig of 14174 bp in length
* 159656 169755: gap of unknown length
* 169756 187711: contig of 17956 bp in length
* 187712 205706: contig of unknown length
* 205707 205807: gap of unknown length
* 205807 228999: contig of 23193 bp in length.
FEATURES
    source
        1. .228999
            Location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="UNK"
                /clone="RP11-1245A23"
                1. .1992
                    /note="assembly_name:Contig10"
                    clone_end:T7
                    vector_side:left"
misc_feature
    /note="assembly_name:Contig11"
    2093..3780
misc_feature
    /note="assembly_name:Contig12"
    3881..6354
misc_feature
    /note="assembly_name:Contig13"
    6455..9843
misc_feature
    /note="assembly_name:Contig14"
    9944..12466
misc_feature
    /note="assembly_name:Contig15"
    12567..15498
misc_feature
    /note="assembly_name:Contig16"
    15599..19057
misc_feature
    /note="assembly_name:Contig17"
    19158..22100
misc_feature
    /note="assembly_name:Contig18"
    22201..25410
misc_feature
    /note="assembly_name:Contig19"
    25511..30906
misc_feature
    /note="assembly_name:Contig20"
    31007..35604
misc_feature
    /note="assembly_name:Contig21"
    35705..39822
misc_feature
    /note="assembly_name:Contig22"
    39923..43181
misc_feature
    /note="assembly_name:Contig23"
    43282..47486
misc_feature
    /note="assembly_name:Contig24"
    47587..52445
misc_feature
    /note="assembly_name:Contig25"
    52546..57196
misc_feature
    /note="assembly_name:Contig26"
    57297..61855
misc_feature
    /note="assembly_name:Contig27"
    61956..66545
misc_feature
    /note="assembly_name:Contig28"
    66646..71594
misc_feature
    /note="assembly_name:Contig29"
    71695..78620
misc_feature
    /note="assembly_name:Contig30"
    78721..85423
misc_feature
    /note="assembly_name:Contig31"
    85524..91961
misc_feature
    /note="assembly_name:Contig32"
    92062..98644
misc_feature
    /note="assembly_name:Contig33"
    98745..105872
misc_feature
    /note="assembly_name:Contig34"
    105973..114931
misc_feature
    /note="assembly_name:Contig35"
    115032..124036

```

```

misc_feature /note="assembly_name:Contig35"
124137..134245
/note="assembly_name:Contig36"
134346..143789
misc_feature /note="assembly_name:Contig37"
143890..153881
/note="assembly_name:Contig38"
153882..159655
misc_feature /note="assembly_name:Contig39"
159756..187711
/note="assembly_name:Contig40"
187812..205706
misc_feature /note="assembly_name:Contig41"
205807..228999
/note="assembly_name:Contig42"
205807..228999
BASP COUNT 71893 a 40635 c 40440 g 72819 t 3212 others
ORIGIN

```

```

Alignment Scores:
Pred. No.: 248 Length: 228999
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.69% Indels: 0
DB: 2 Gaps: 0

```

US-09-092-297-17 (1-117) x AC083885 (1-228999)

QY 16 AlavAlphealpheCysIleSer 24

DB 145812 GCTGCTTTCACATATCTCATTTCA 145838

```

RESULT 9
LOCUS G00569 329 bp DNA linear STS 28-FEB-1995
DEFINITION Fruit fly STS Dm0281 clone 1(3)03576, sequence tagged site.
ACCESSION G00569
VERSION G00569.1 GI:683973
KEYWORDS STS: STS sequence; primer; sequence tagged site.
SOURCE fruit fly vector-P2[ry, lacZ] STS's whose source is listed as a
spreading P-element screen are obtained from plasmid rescues of
P-element stocks from a P-element lethal screen (see Karpen and
Spradling, Genetics 132: 737). The number in parentheses in the
clone name designates the chromosome the P-element is on. Genetic
characterization of these P-element insertions was done in A.
Spradling's lab; these P-element insertions each inactivate a
different essential genetic locus. The chromosomal location of the
P-element insertion was mapped by T. Laverly in G. Rubin's lab.
STS's were generated by plasmid rescue and sequencing of the
genomic region immediately flanking the P-element insertion site,
by the W. Kimmerly, C. Martin, and M. Palazzolo lab at LBL.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 329)
Rubin.G.
Drosophila STS
Unpublished

```

```

ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact:
Berkeley Drosophila Genome Project
Primer A: CAGGTGACCATGATGCAATA
Primer B: AATTGATTTTGGATCAACA
STS size: 153
PCR Profile:
Annealing: 55 degrees C
PCR Cycles: 30
Protocol:
Template: P1 Library Pools
Primer: 1 uM each
dNTPs: 250 uM each

```

```
* 100395 100494: gap of unknown length
* 100435 102562: contig of 2068 bp in length
* 102563 102662: gap of unknown length
* 102663 105607: contig of 2945 bp in length
* 105608 105707: gap of unknown length
* 105708 108535: contig of 2828 bp in length
* 108536 108635: gap of unknown length
* 108636 111168: contig of 2533 bp in length
* 111169 111269: gap of unknown length
* 111269 114040: contig of 2772 bp in length
* 114041 114140: gap of unknown length
* 114141 117073: contig of 2833 bp in length
* 117074 117173: gap of unknown length
* 117174 119836: contig of 2663 bp in length
* 119837 119936: gap of unknown length
* 123089 123089: contig of 3153 bp in length
* 123090 123189: gap of unknown length
* 123190 125696: contig of 2507 bp in length
* 125697 125796: gap of unknown length
* 125797 128622: contig of 2826 bp in length
* 128623 128722: gap of unknown length
* 131944 131944: contig of 3222 bp in length
* 131945 132044: gap of unknown length
* 132045 135330: contig of 3286 bp in length
* 135331 135430: gap of unknown length
* 135431 138631: contig of 3201 bp in length
* 138632 138731: gap of unknown length
* 138732 143021: contig of 4290 bp in length
```

Alignment Scores:

```
Pred. No.: 198 Length: 180563
Percent: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.69% Indels: 0
DB: 2 Gaps: 0
```

US-09-092-297-17 (1-117) x AC034170 (1-180563)

QY 21 PhycysilesesArgleuLeucyCysSer 29

Db 90512 TTTTGCATTTCTCTCTCTCTCTCTCA 90486

RESULT 8

AC083885 228999 bp DNA linear HTG 30-OCF-2000

LOCUS Homo sapiens chromosome UNK clone RP11-1245A23, WORKING DRAFT

AC083885 SEQUENCE 33 unordered pieces.

AC083885.3 GI:11038707

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 228999)

Waterston,R.H.

The sequence of Homo sapiens clone

Unpublished

2 (bases 1 to 228999)

Waterston,R.H.

Direct Submission

Submitted (04-OCT-2000) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

On Oct 30, 2000 this sequence version replaced gi:11024957.

COMMENT

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu/gsc/index.shtml

Project Information

Center project name: H_NH1245A23

----- Summary Statistics -----

```
Sequencing vector: M13: 100%
Sequencing vector: plasmid: 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 214460 bases at least Q40
Consensus quality: 219291 bases at least Q30
Consensus quality: 221570 bases at least Q20
Insert size: 214000; agarose-fp
Insert size: 225799; sum-of-contigs
Quality coverage: 4.55 in Q20 bases; agarose-fp
Quality coverage: 4.35 in Q20 bases; sum-of-contigs
```

* NOTE: This is a 'working draft' sequence. It currently consists of 33 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 1992: contig of 1992 bp in length
1993 2092: gap of unknown length
2093 3780: contig of 1688 bp in length
3781 3880: gap of unknown length
3881 6354: contig of 2474 bp in length
6355 6454: gap of unknown length
6455 9843: contig of 3389 bp in length
9844 12465: gap of unknown length
12466 12565: gap of unknown length
12567 15498: contig of 2932 bp in length
15499 15598: gap of unknown length
15599 19057: contig of 3459 bp in length
19058 19157: gap of unknown length
19158 22100: contig of 2943 bp in length
22101 25410: gap of unknown length
25411 25510: contig of 3210 bp in length
25511 30906: gap of unknown length
30907 31006: contig of 5396 bp in length
31007 35604: gap of unknown length
35605 35704: contig of 4598 bp in length
35705 39822: gap of unknown length
39823 39922: contig of 4118 bp in length
39923 43181: gap of unknown length
43182 43282: contig of 3259 bp in length
43283 47486: gap of unknown length
47487 47586: contig of 4205 bp in length
47587 52445: gap of unknown length
52446 52545: contig of 4859 bp in length
52546 57196: gap of unknown length
57197 57296: contig of 4651 bp in length
57297 61855: gap of unknown length
61856 61955: contig of 4559 bp in length
61956 66545: gap of unknown length
66546 66645: contig of 4590 bp in length
66646 71594: gap of unknown length
71595 71694: contig of 4949 bp in length
71695 78620: gap of unknown length
78621 78720: contig of 6926 bp in length
78721 85424: gap of unknown length
85425 85523: contig of 6703 bp in length
85524 91961: gap of unknown length
91962 92061: contig of 6438 bp in length
92062 98644: gap of unknown length
98645 98744: contig of 6583 bp in length
98745 105872: gap of unknown length
105873 105972: contig of 7128 bp in length
105973 114931: gap of unknown length
114932 115031: contig of 8959 bp in length
115032 124036: contig of 9005 bp in length
```

----- Project Information -----
Center project name: H_NH0674UD5
* NOTE: This is a 'working draft' sequence. It currently
* consists of 89 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 699: contig of 699 bp in length
700 799: gap of unknown length
800 1512: contig of 713 bp in length
1513 1612: gap of unknown length
1613 2229: contig of 617 bp in length
2230 2329: gap of unknown length
2330 2944: contig of 615 bp in length
2945 3044: gap of unknown length
3045 3953: contig of 909 bp in length
3954 4053: gap of unknown length
4054 4783: contig of 730 bp in length
4784 4883: gap of unknown length
4884 5571: contig of 688 bp in length
5572 5671: gap of unknown length
5672 6377: contig of 705 bp in length
6377 6476: gap of unknown length
6477 7207: contig of 731 bp in length
7208 7307: gap of unknown length
7308 7613: contig of 306 bp in length
7614 7713: gap of unknown length
7714 8902: contig of 1189 bp in length
8903 9002: gap of unknown length
9003 9757: contig of 755 bp in length
9758 9857: gap of unknown length
9858 10343: contig of 486 bp in length
10344 10443: gap of unknown length
10444 11201: contig of 758 bp in length
11202 11301: gap of unknown length
11302 12335: contig of 1034 bp in length
12336 12435: gap of unknown length
12436 13294: contig of 859 bp in length
13295 13394: gap of unknown length
13395 14096: contig of 702 bp in length
14097 14196: gap of unknown length
14197 15252: contig of 1056 bp in length
15253 15352: gap of unknown length
15353 16666: contig of 1314 bp in length
16667 16766: gap of unknown length
16767 17422: contig of 656 bp in length
17423 17522: gap of unknown length
17523 18416: contig of 894 bp in length
18417 18516: gap of unknown length
18517 19058: contig of 542 bp in length
19059 19158: gap of unknown length
19159 20107: contig of 949 bp in length
20108 20207: gap of unknown length
20208 21393: contig of 1185 bp in length
21393 21493: gap of unknown length
21493 22510: contig of 1018 bp in length
22511 22610: gap of unknown length
22610 23648: contig of 1038 bp in length
23649 23748: gap of unknown length
23749 24975: contig of 1227 bp in length
24976 25075: gap of unknown length
25076 26075: gap of 1000 bp in length
26076 26175: gap of unknown length
26176 27416: contig of 1241 bp in length
27417 27516: gap of unknown length
27517 28529: contig of 1013 bp in length
28530 28629: gap of unknown length
28630 29583: contig of 954 bp in length
29584 29683: gap of unknown length
29684 30741: contig of 1058 bp in length
30742 30841: gap of unknown length
30842 31989: contig of 1148 bp in length
31989 32089: gap of unknown length
32090 33154: contig of 1065 bp in length
33155 33255: gap of unknown length
33255 34457: contig of 1203 bp in length
34458 34557: gap of unknown length
34558 35986: contig of 1429 bp in length
35987 36086: gap of unknown length
36087 37925: contig of 1839 bp in length
37926 38025: gap of unknown length
38026 39294: contig of 1269 bp in length
39295 39394: gap of unknown length
39395 40532: contig of 1138 bp in length
40533 40633: gap of unknown length
40633 42129: contig of 1497 bp in length
42130 42229: gap of unknown length
42230 43855: contig of 1626 bp in length
43856 43955: gap of unknown length
43956 45742: contig of 1787 bp in length
45743 45842: gap of unknown length
45843 46799: contig of 957 bp in length
46799 46800: gap of unknown length
46800 48309: contig of 1410 bp in length
48310 48409: gap of unknown length
48410 49768: contig of 1359 bp in length
49769 49868: gap of unknown length
49869 51673: contig of 1804 bp in length
51673 51773: gap of unknown length
51773 53070: contig of 1298 bp in length
53071 53171: gap of unknown length
53171 55890: contig of 2720 bp in length
55891 55990: gap of unknown length
55991 58540: contig of 2550 bp in length
58541 58640: gap of unknown length
58641 60742: contig of 2102 bp in length
60743 60842: gap of unknown length
60843 62553: contig of 1710 bp in length
62553 62653: gap of unknown length
62653 65163: contig of 2511 bp in length
65164 65263: gap of unknown length
65264 66984: contig of 1721 bp in length
66985 67084: gap of unknown length
67085 68550: contig of 1466 bp in length
68551 68650: gap of unknown length
68651 70725: contig of 2075 bp in length
70726 70825: gap of unknown length
70826 72615: contig of 1790 bp in length
72616 72715: gap of unknown length
72716 74764: contig of 2049 bp in length
74765 74864: gap of unknown length
74865 77074: contig of 2210 bp in length
77075 77175: gap of unknown length
77175 79340: contig of 2166 bp in length
79341 79440: gap of unknown length
79441 80990: contig of 1550 bp in length
80991 81090: gap of unknown length
81091 82975: contig of 1885 bp in length
82976 83075: gap of unknown length
83076 84926: contig of 1851 bp in length
84927 85026: gap of unknown length
85027 87554: contig of 2528 bp in length
87555 87654: gap of unknown length
87655 90562: contig of 2908 bp in length
90563 90662: gap of unknown length
90663 92286: contig of 1624 bp in length
92287 92386: gap of unknown length
92387 95351: contig of 2965 bp in length
95352 95451: gap of unknown length
95452 97620: contig of 2169 bp in length
97621 97720: gap of unknown length
97721 100394: contig of 2674 bp in length

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: http://gc.bcm.tmc.edu:8088/quality_info/genbank.annotation.html.

QUALSTART-REPORT-----

Summary Statistics -----

Contig length: 172246
 Phrap values in estimate: 171130
 Average error rate (BCM-Phrap estimate): 2.13165e-05
 Fraction of Phrap values less than 40 : 0.00600713
 Number of consensus changing edits: 13
 Number of N's in consensus : 0

Position	Original+Context	Edited+Context
67598	tcttagacag(n)aatgacata	tcttagacag(c)aatgacata
99512	ctcttgaag(n)atattatg	ctcttgaag(c)atattatg
103635	ggcatggg(n)nacagggccc	ggcatggg(g)(c)aacagggccc
103636	ggcatggg(n)nacagggccc	ggcatggg(g)(c)aacagggccc
103660	aaagactgga(n)ccaggaatg	aaagactgga(a)ccaggaatg
106704	taataacct(a)ataaaanta	taataacct(t)ataaaanta
106705	taataacct(a)ataaaanta	taataacct(t)ataaaanta
106713	ctctgagtc(c)tagcttagt	ctctgagtc(c)tagcttagt
117425	cttaactat(n)aaagcttag	cttaactat(c)aaagcttag
160258	acactcgtc(n)atcgcctg	acactcgtc(a)atcgcctg
169390	caagactgaa(a)tc	caagactgaa(t)tc
172244	agactgaaat(t)	agactgaaat(c)
172246		

----- Distribution of Quality < 40 Bases -----

	5	10	15	20	25	30	35	40
10001								
9001								
8001								
7001								
6001								
5001								
4001								
3001								
2001								
1001								
01	*	*	*	*	*	*	*	*

Phrap Value Range

Version: 1.01 gxf.

Location/Qualifiers

source
 1. 172246
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="12"
 /clone="RP11-900F13"
 /complement(7..342)
 /rpt_family="MLT2CA"
 /complement(674..798)
 /rpt_family="MIR"
 /complement(799..835)
 /rpt_family="LIPAS"
 /complement(836..852)
 /rpt_family="MIR"
 913..957
 /rpt_family="(TC)n"
 1986..2007
 repeat_region

repeat_region	/rpt_family="AT-rich"	2317..2346	length: 172246
repeat_region	/rpt_family="(TTTG)n"	complement(2347..2628)	Score: 9.00
repeat_region	/rpt_family="AluY"	complement(5766..6002)	Percent Similarity: 100.00%
repeat_region	/rpt_family="MSTA"	6004..6045	Best Local Similarity: 100.00%
repeat_region	/rpt_family="(TTTTG)n"	complement(6046..7170)	Query Match: 9
repeat_region	/rpt_family="LHS"	7168..7228	
repeat_region	/rpt_family="LHS"	complement(7229..7423)	
repeat_region	/rpt_family="MSTA"	7642..7916	
repeat_region	/rpt_family="L1"	8451..8649	
repeat_region	/rpt_family="AluO"	complement(8659..8759)	
repeat_region	/rpt_family="MER5A"	9713..10458	
repeat_region	/rpt_family="L2"	complement(10863..11134)	
repeat_region	/rpt_family="L2"	11942..11974	
repeat_region	/rpt_family="AT-rich"	complement(11340..13408)	
repeat_region	/rpt_family="MER63A"	14596..14950	

Alignment Scores:

Pred. No.: 189
 Score: 9.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 7.69%
 DB: 9
 Gaps: 0

US-09-092-297-17 (1-117) x AC024941 (1-172246)

OY 21 PhcysileSerArgLeuLeucysSer 29

Db 130738 TTTTCATTCTCGCTCTCTGCTCA 130712

RESULT 7

AC034170/c 180563 bp DNA linear HTG 20-APR-2000

LOCUS Homo sapiens chromosome 12 clone RP11-674J5, *** SEQUENCING IN

AC034170 AC034170.3 GI:7622528

ACCESSION AC034170.3 GI:7622528

VERSION HTG; HTGS_PHASE1.

KEYWORDS HTG; HTGS_PHASE1.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 180563)

AUTHORS Waterston,R.H.

TITLE The sequence of Homo sapiens clone

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 180563)

AUTHORS Waterston,R.H.

TITLE Direct Submission

JOURNAL Submitted (04-APR-2000) Genome Sequencing Center, Washington

COMMENT MO 63108, USA

On Apr 20, 2000 this sequence version replaced gl:7579885.

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

```

repeat_region      /rpt_family="L1"
                    83744..83767
                    /rpt_family="L1"
repeat_region      83784..83813
                    /rpt_family="L1"
repeat_region      84400..84452
                    /rpt_family="L1"
repeat_region      84599..84655
                    /rpt_family="L1"
repeat_region      complement(85751..85835)
                    /rpt_family="L1"
repeat_region      85770..85847
                    /rpt_family="L1"
repeat_region      89152..89193
                    /rpt_family="L1"
repeat_region      93896..93929
                    /rpt_family="L1"
repeat_region      complement(95512..95708)
                    /rpt_family="ALU"
repeat_region      97302..97346
                    /rpt_family="L1"
repeat_region      98231..98289
                    /rpt_family="L1"
repeat_region      98500..98555
                    /rpt_family="L1"
repeat_region      complement(99131..99423)
                    /rpt_family="ALU"
repeat_region      107834..107888
                    /rpt_family="L1"
repeat_region      108592..108880
                    /rpt_family="L1"
repeat_region      108881..108923
                    /rpt_family="ALU"

```

Alignment Scores:

Pred. No.:	Score:	Length:
Percent Similarity:	175	158537
Best Local Similarity:	9.00	Matches:
Query Match:	100.008	Conservative:
	7.698	Mismatches:
		Indels:
		Gaps:

US-09-092-297-17 (1-117) x HSAC002069 (1-158537)

Qy 16 Alvalphaatllephcyslleser 24

Db 15848 GCTGCTTTCACATATTCGACATTCA 15822

RESULT 6

AC024941/c

LOCUS

DEFINITION

AC024941

AC024941.30

HTG

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AC024941 172246 bp DNA linear PRI 11-JUL-2001
 Homo sapiens 12 BAC RP11-900F13 (Roswell Park Cancer Institute
 Human BAC Library) complete sequence.
 AC024941.30 GI:14669928
 HTG
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 172246)
 Muzny,D.M., Adams,C.C., Adio-oduola,B., All-osman,F.R., Allen,C.,
 Alstbrooks,S.L., Amaralungu,H.C., Are,J.R., Banks,T., Barbarella,J.,
 Benton,J., Blinaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
 Bowles,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
 Chen,Z., Chiu,D., Chowdhury,I., Christopoulos,C., Cleveland,C.D.,
 Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
 Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
 Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,
 Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisti,A., Gao,J.,

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Garcla,A., Garner,T., Garza,N., Gill,R., Girelli,J.H., Guevara,M.,
 Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,
 Hart,M., Haylak,P., Hawes,A., Hernandez,J., Hernandez,O.,
 Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homai,F.,
 Howard,S., Huber,J., Hulik,S., Hume,J., Ioshikhes,I., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Joliver,S., Joudah,S.,
 Karisson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,
 Lewis,L., Li,J., Li,Z., Lichtenberg,O., Lieu,C., Liu,J., Liu,W.,
 Louisedge,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
 Ma,J., Maheshwari,M., Mapa,P., Marondel,I., Martin,R.,
 Martindale,A., Martinez,E., Massey,E., Maxwell,E., McLeod,M.P.,
 Meador,M., Mei,G., Merscher,S., Metzger,M., Miller,A., Miner,G.,
 Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,
 Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,
 Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nokenko,S.,
 Oguh,M., Okunnu,G., Oragunye,N., Oriado,R., Pace,A., Payton,B.,
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
 Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
 Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shm,C.,
 Shooshitari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A.,
 Stanley,H., Stone,H., Sutton,A., Svatok,A., Tabot,P., Tamerisa,A.,
 Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telitod,B.,
 Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,
 Vanson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,
 Washington,C., Watlington,S., Williams,G., Williamson,A.,
 Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,
 Zorilla,S., Kuchelapatti,R. and Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 172246)
 Morley,K.C.
 Direct Submission
 Submitted (03-MAR-2000) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 172246)
 Morley,K.C.
 Direct Submission
 Submitted (11-JUL-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 11, 2001 this sequence version replaced g1:13492990.
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
 gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
 entire insert of this clone. Overlapping regions of clones are only
 sequenced and submitted once, so the sequence for the remainder of
 the insert may be found in the record for the adjacent clones.
 Overlapping clones are noted at the beginning and end of the
 Features listing.

ANNOTATION OF FEATURES:

STRs are identified using ePCR (Genome Res. 7:541-550) searches
 of a local database that includes entries from dbSTS, GDB, and
 local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
 unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST
 (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
 EST and cDNA sequences. Genes demonstrate at least two exons
 flanked by consensus splice sites that maintained sequence
 continuity across the splice junctions. Sequences that are not
 identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
 standard of double strand coverage with a minimum of 2 clones and 2
 reads with no ambiguities or 2 chemistries with a minimum of 2
 clones and 3 reads with no ambiguities. If the sequence quality for
 a region does not meet this standard, it will be indicated in the
 annotation as Low Coverage.

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7> or send an E-mail to egreen@nhgri.nih.gov

SOURCE INFORMATION:

This clone is from a release of the human BAC library. The library contains cloned DNA from a human sperm. For references see: Shizuya et al., Proc. Natl. Acad. Sci. 89:8794-8797 (1992); Kim et al., Genomics 34:213-218 (1996).

VECTOR: pBEO

Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

Actual start of this clone is at base position 1 of H_RG326K09; actual end is at 158537 of H_RG326K09. This clone lies in an unanchored cluster of unknown orientation.

This clone contains SWS3233 (NID:g1113712).

FEATURES:

source

location/Qualifiers

1..158537

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="7"

/map="7q21"

/clone="RG326K09"

/clone_1lb="CTB-HS-A"

276..1999

/rpt_family="L1"

complement(580..998)

3260..3282

/rpt_family="L1"

4062..4178

/rpt_family="L1"

4193..4376

/rpt_family="L1"

4553..4725

/rpt_family="L1"

4767..4808

/rpt_family="L1"

5097..6621

/rpt_family="L1"

complement(5236..5642)

6764..6910

/rpt_family="L1"

6922..6974

/rpt_family="L1"

complement(6997..7223)

7606..7712

/rpt_family="ALU"

7866..7936

/rpt_family="L1"

11439..11461

/rpt_family="L1"

repeat_region

11641..11715

repeat_region

12310..12397

repeat_region

12589..12629

repeat_region

13203..13236

repeat_region

13834..14031

repeat_region

16674..17204

repeat_region

complement(17532..17751)

repeat_region

complement(19040..19066)

misc_feature

20125..21945

misc_feature

(note="Probable pseudogene. Similar to hSRPalpha U28386 (NID:g899538).")

misc_feature

20132..20288

misc_feature

(note="Similar to human EST W56840 (NID:g1358716) zc97d02.r1")

repeat_region

complement(26748..28317)

repeat_region

29378..35406

repeat_region

complement(33995..34413)

repeat_region

complement(35884..35905)

repeat_region

complement(35910..36194)

misc_feature

40798..45213

repeat_region

(note="Probable retroviral pseudogene. Similar to POL_GALV POL POLYPROTEIN P21414 (NID:g130602).")

repeat_region

complement(46549..47052)

repeat_region

complement(47102..47157)

repeat_region

47521..50154

repeat_region

51874..52163

repeat_region

53038..53063

repeat_region

complement(56184..56476)

repeat_region

56478..56736

repeat_region

59292..60851

repeat_region

61080..61206

repeat_region

complement(65139..65325)

repeat_region

75780..76108

repeat_region

78275..78293

repeat_region

79752..79781

repeat_region

80350..80411

repeat_region

80412..80438

repeat_region

80761..83307

repeat_region

83308..83601

repeat_region

83602..83652

```

repeat_region /rpt_family="Alusq"
10273. .10347
repeat_region /rpt_family="FLAM"
10691. .10818
repeat_region /rpt_family="Aluub"
complement(10999. .11283)
repeat_region /rpt_family="AluJo"
complement(11302. .11517)
repeat_region /rpt_family="LIMC/D"
complement(11536. .11661)
repeat_region /rpt_family="FLAM_C"
complement(11667. .11798)
repeat_region /rpt_family="LIMC/D"
complement(11833. .12138)
repeat_region /rpt_family="Alusq"
complement(12158. .12478)
repeat_region /rpt_family="LIMC/D"
12479. .12642
repeat_region /rpt_family="FRAM"
complement(12655. .12795)
repeat_region /rpt_family="LIMC/D"
complement(12806. .13106)
repeat_region /rpt_family="Alusx"
complement(13356. .13481)
repeat_region /rpt_family="(GAA)n"
complement(13482. .13694)
repeat_region /rpt_family="Aluub"
complement(13709. .14037)
repeat_region /rpt_family="L1"
complement(14009. .14105)
repeat_region /rpt_family="LIMC/D"
complement(14115. .14415)
repeat_region /rpt_family="Aluy"
complement(14417. .14849)
repeat_region /rpt_family="LIMC/D"
complement(14859. .15153)
repeat_region /rpt_family="Alusx"
15226. .15492
repeat_region /rpt_family="AluJo"
complement(16257. .16340)
misc_feature /note="Predicted exon, program: graal2exons_human_1.3,
16362. .16666
frame: 0, quality: good, score: 63.000"
repeat_region /rpt_family="Alusq"
complement(16991. .17060)
repeat_region /rpt_family="THE1C"
17482. .17781
repeat_region /rpt_family="AluJo"
17782. .17870
repeat_region /rpt_family="(GAA)n"
17968. .18273
repeat_region /rpt_family="Alusq"
complement(18401. .18817)
repeat_region /rpt_family="LTR7"
complement(18818. .19193)
repeat_region /rpt_family="HERVH"
complement(19194. .19412)
repeat_region /rpt_family="HERVH"
complement(19414. .20028)
repeat_region /rpt_family="HERVH"
complement(20031. .23647)
repeat_region /rpt_family="HERVH"
complement(23651. .24066)
repeat_region /rpt_family="LTR7"
complement(24403. .24714)
repeat_region /rpt_family="MER74"
24789. .25080
repeat_region /rpt_family="Alusx"
25161. .25207
repeat_region /rpt_family="LIMD3"
25224. .25509
repeat_region /rpt_family="Aluub"
25605. .25901

```

```

repeat_region /rpt_family="Alusq"
25902. .26069
repeat_region /rpt_family="Alusq/x"
complement(26224. .26256)
repeat_region /rpt_family="(TAA)n"
complement(26257. .26545)
repeat_region /rpt_family="AluJo"
26573. .26876
repeat_region /rpt_family="AluJo"
complement(27125. .27422)
repeat_region /rpt_family="Alusx"
27866. .28173
repeat_region /rpt_family="Alusx"
28181. .28279
repeat_region /rpt_family="AluJ/FLAM"
complement(28359. .28522)
repeat_region /rpt_family="Alusq"
complement(28526. .28705)
repeat_region /rpt_family="AlusC"
complement(28733. .28872)
repeat_region /rpt_family="Aluub"
28874. .29081
misc_feature /rpt_family="AluJo"
29463. .29583
/rnote="Predicted exon, program: graal2exons_human_1.3,
frame: 2, quality: good, score: 71.000"
complement(29585. .29888)
repeat_region /rpt_family="Alusq"
30336. .30634
repeat_region /rpt_family="Alusq"
30646. .30779
repeat_region /rpt_family="Aluub"
complement(30830. .31040)
/rpt_family="MIR"

```

```

Alignment Scores:
Pred. No.: 162
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 7.69%
DB: 9
Gaps: 0

```

US-09-092-297-17 (1-117) x AC007193 (1-146180)

QY 25 ArgLeuGlySerHisGlyAlaPro 33

DB 6042 AGGCTCTCTGCTCACACGAGCCCA 6016

RESULT 5
HSAC002069/c 158537 bp DNA linear PRI 09-MAY-1997

LOCUS Human BAC clone RG326K09 from 7q21, complete sequence.

ACCESSION AC002069

VERSION AC002069.1 GI:2076718

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 158537)

AUTHORS Geisel,C., Beck,C., Smith,A. and Twyman,B.

TITLE The sequence of H. sapiens BAC clone RG326K09

JOURNALS Unpublished (1997)

REFERENCE 2 (bases 1 to 158537)

AUTHORS Waterston,R.

TITLE Direct Submission

JOURNALS Submitted (09-MAY-1997)

COMMENT Genome Sequencing Center

Department of Genetics, Washington University

St. Louis, MO 63108, USA

http://genome.wustl.edu/gsc

e-mail: saplens@wustl.wustl.edu

```

/note="Predicted exon, program: grail2exons_human_1.3,
frame: 2, quality: good, score: 50.000"
repeat_region 29332..29526
/rpt_family="MERAD"
repeat_region 29527..29826
/rpt_family="AluSg"
repeat_region 29827..29870
/rpt_family="MERAD"
repeat_region 29870..30491
/rpt_family="MERAD"
repeat_region 30523..30611
/rpt_family="(GA)n"
repeat_region 31175..31465
/rpt_family="AluSg"
repeat_region complement(31707..32005)
/rpt_family="AluSg"
repeat_region complement(32050..32614)
/rpt_family="MLT2D"
repeat_region 32619..32691
/rpt_family="LTR8"

```

```

Alignment Scores:
Pred. No.: 1.64e-82 Length: 107573
Score: 82.00 Matches: 82
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 70.09% Indels: 0
DB: 9 Gaps: 0

```

US-09-092-297-17 (1-117) x AC006262 (1-107573)

```

QY 33 ProValAlaProMetThrProTyrLeuMetLeuGlyGlnProHisLysArgCysGlyasp 52
Db 81800 CCACTGGCCCCACGACCTTACTGATGCTGTGTCACAGCACACAGATGATGGGAC 81859
QY 53 LysPheTyrAspProLeuGlnHisCysGlyTyrAspAlaValAlaProLeuAlaArg 72
Db 81860 AAGTTCACGACCCCTGACACACTGTTGTAATGATGATGCCGTCGTCCTGGCAGG 81919
QY 73 ThcGlnThrCysGlyAsnCysTthrPheArgValCysPheGlnGlnCysCysProTrrThr 92
Db 81920 ACCAGACAGTGTGGAACCTGACCTTGCAGTCTGCTTGAGCAGTGTGTCCTGGCAGC 81979
QY 93 PheMetValLysLeuIleAsnGlnAsnCysAspSerAlaArgThrSerAspAspArgLeu 112
Db 81980 TTCATGTTGAGATGATTAACACAGACTGCATCAGCCGACCGACTCGATGACAGCCTT 82039
QY 113 CysArg 114
Db 82040 TGTCCG 82045

```

```

RESULT 4
AC007193 146180 bp DNA linear PRI 17-JUN-1999
LOCUS Homo sapiens chromosome 19, BAC 82621 (CIT-B-139a18), complete
DEFINITION
AC007193
AC007193
AC007193.1 GI:4558635
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 146180)
Lamerdin,J.E., McCready,P.M., Skowronski,E., Viswanathan,V.,
Burkhardt-Schultz,K.J., Gordon,L., Dias,J., Ramirez,M.,
Stillwagen,S., Phan,H., Velasco,N., Do,L., Regala,W., Terry,A.,
Garnes,V., Dangnanan,L., Erlar,A., Christensen,M., Georgescu,A.,
Avila,J., Liu,S., Altix,C., Andreise,T., Trankheim,M.,
Amico-Keller,G., Coefield,J., Duarte,S., Lucas,S., Bruce,R.,
Thomas,P., Quan,G., Krommiller,B., Arellano,A., Sanders,C., Ow,D.,
Nolan,M., Trong,S., Kobayashi,A., Olsen,A.S. and Carraro,A.V.
Sequence analysis of a 1.9 Mb region in 19q13.2 between APOE and

```

```

JOURNAL D19S412
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 146180)
TITLE Lamerdin,J.E.
JOURNAL Direct Submission
COMMENT Submitted (02-APR-1999) Joint Genome Institute, Lawrence Livermore
National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
Map and sequence oriented from 9 centromere to telomere. BAC 82621
(CIT-B-139a18) overlaps BAC 264576 (CIT-B-297a14) to the left from
bases 1 to 27,487 of this accession, and separated from BAC 89981
(CIT-B-158d10) on the right by a gap of approximately 17 to 23 kb.
Additional chromosome 19 map and sequence information may be
obtained at: http://www.bio.llnl.gov/dbp/genome/genome.html.

```

FEATURES

```

source
1..146180
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2 between APOE and D19S412"
/clone="BC82621"
/cell_line="987SK"
/cell_type="fibroblast"
/note="LNL clone name: BC82621 BAC library obtained from
Research Genetics."
1..101
/rpt_family="L1"
511..677
/rpt_family="L1"
1018..1158
/rpt_family="L1"
1159..1453
/rpt_family="AluSx"
1459..1529
/rpt_family="L1"
complement(1857..1974)
/rpt_family="FLAM_A"
2115..2375
/rpt_family="L1MB4"
2478..2549
/rpt_family="LINE2"
complement(2629..2921)
/rpt_family="AluB"
3355..3655
/rpt_family="AluSg"
3673..3939
/rpt_family="L1"
complement(4005..4264)
/rpt_family="AluSx"
complement(4269..4568)
/rpt_family="AluY"
5064..5122
/note="Predicted exon, program: grail2exons_human_1.3,
frame: 2, quality: excellent, score: 81.000"
complement(5169..5468)
/rpt_family="AluY"
5626..5886
/note="Predicted exon, program: grail2exons_human_1.3,
frame: 0, quality: good, score: 52.000"
misc_feature
repeat_region
7398..7444
/rpt_family="Alu"
7480..7595
/rpt_family="(TA)n"
7596..7812
/rpt_family="AluB"
7900..8191
/rpt_family="AluSx"
complement(8575..8597)
/rpt_family="AT-rich"
complement(8598..8878)
/rpt_family="AluSx"
9477..9766
/rpt_family="AluSx"
9997..10258
repeat_region

```

overlaps BC282485 (CFC-344H19, AC007785) to the left from bases 1 to 20,296 of this accession, and overlaps BC82621 (CTR-139A18, AC007193) to the right from bases 106,317 to 107,573 of this accession. This sequence does not represent the entire insert of BC264576. Additional chromosome 19 map and sequence information may be obtained at: <http://www-bio.lm1.gov/chrp/genome/genome.html>.

FEATURES

source

1. 107573

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="19"

/map="19q13.2 between APOE and D19S412"

/clone="CFC-297N14"

/cell_line="978SK"

/cell_type="fibroblast"

/note="LNL clone name: BC264576 BAC library obtained from Research Genetics."

354. .567

/note="DDS similarity to AA456254 zx99d06.r1 Soares NHHMPUS1 Homo sapiens CDNA clone 811883 5'; (3. .215); 100%

identity."

repeat_region complement(669. .800)

/rpt_family="GA)n"

repeat_region complement(829. .953)

/rpt_family="(GGGA)n"

repeat_region complement(1035. .2017)

/rpt_family="LIMB7"

repeat_region complement(2014. .2202)

/rpt_family="LIMA7"

repeat_region complement(2206. .2495)

/rpt_family="AluJo"

repeat_region complement(2497. .2728)

/rpt_family="LIMA7"

repeat_region complement(2730. .3133)

/rpt_family="LIMB7"

repeat_region complement(3158. .3447)

/rpt_family="AluJo"

repeat_region complement(3448. .3803)

/rpt_family="LIMB7"

repeat_region complement(3805. .4108)

/rpt_family="AluJo"

repeat_region complement(4110. .4801)

/rpt_family="LIMB7"

repeat_region complement(4826. .5065)

/rpt_family="LIM1"

repeat_region complement(5338. .5369)

/rpt_family="(TA)n"

repeat_region complement(5370. .5708)

/rpt_family="LIPAB"

repeat_region complement(5974. .6277)

/rpt_family="AluY"

repeat_region complement(7059. .7080)

/rpt_family="AT-rich"

repeat_region complement(7150. .7485)

/rpt_family="LI"

repeat_region complement(7527. .7646)

/rpt_family="FLAM-C"

repeat_region complement(7661. .7916)

/rpt_family="LI"

repeat_region complement(9725. .9725)

/rpt_family="LI"

repeat_region complement(9736. .10038)

/rpt_family="AluSg"

repeat_region complement(10866. .10964)

/note="predicted exon, program: grail2exons_human_1.3, frame: 1, quality: excellent, score: 75.000"

repeat_region 11326. .11517

/rpt_family="AluJo"

repeat_region 11945. .11998

/note="DDS similarity to AA456254 zx99d06.r1 Soares NHHMPUS1 Homo sapiens CDNA clone 811883 5'; (216. .269); 100% identity."

misc_feature 12144. .12411

/note="DDS similarity to overlapping ESTs: (12144. .12398) AA456254 zx99d06.r1 Soares NHHMPUS1 Homo sapiens CDNA clone 811883 5'; (270. .523); 99% identity.-(12182. .12411) A1081371 ox76g10.x1 Soares NHHMPUS1 Homo sapiens CDNA clone IMAGE:1662306 3'; (454. .225); 100% identity.-(12173. .12411) AA456442 zx99d06.s1 Soares NHHMPUS1 Homo sapiens CDNA clone 811883 3'; (455. .217); 100% identity.-(12144. .12411) AA760247 vv74e06.r1 StrataGene mouse skin (#937313) Mus musculus CDNA clone 1228162 5'; (119. .380); 69% identity."

/note="predicted exon, program: grail2exons_human_1.3, frame: 2, quality: excellent, score: 90.000"

misc_feature 12600. .12818

/note="DDS similarity to overlapping ESTs: (12603. .12818) AA456442 zx99d06.s1 Soares NHHMPUS1 Homo sapiens CDNA clone 811883 3'; (216. .1); 99%

identity.-(12600. .12823) A1081371 ox76g10.x1 Soares NHHMPUS1 Homo sapiens CDNA clone IMAGE:1662306 3'; (224. .1); 98% identity.-(12600. .12656) AA760247 vv74e06.r1 StrataGene mouse skin (#937313) Mus musculus CDNA clone 1228162 5'; (381. .432); 76% identity."

complement(12967. .16062)

/rpt_family="LIP5"

repeat_region complement(14029. .14445)

/note="predicted exon, program: grail2exons_human_1.3, frame: 0, quality: good, score: 58.000"

misc_feature complement(17695. .18620)

/rpt_family="MER118"

repeat_region complement(18619. .18767)

/rpt_family="MER11A"

repeat_region complement(19519. .19519)

/rpt_family="MLT1B"

repeat_region complement(19592. .19754)

/rpt_family="AluJo"

repeat_region complement(19767. .19870)

/rpt_family="MER65A"

repeat_region complement(20427. .20481)

/rpt_family="(CA)n"

repeat_region complement(20546. .20751)

/rpt_family="AluJo"

repeat_region complement(21128. .21397)

/rpt_family="AluSx"

repeat_region complement(21768. .22798)

/rpt_family="MER11A"

repeat_region complement(23090. .23194)

/note="predicted exon, program: grail2exons_human_1.3, frame: 2, quality: good, score: 50.000"

misc_feature complement(24023. .24141)

/rpt_family="FLAM_A"

repeat_region complement(24258. .24429)

/rpt_family="AluJo"

repeat_region complement(24910. .25309)

/rpt_family="MSTB"

repeat_region complement(25329. .25464)

/rpt_family="HBRVL"

repeat_region complement(25732. .25881)

/rpt_family="MER45"

repeat_region complement(26586. .26613)

/rpt_family="AT-rich"

repeat_region complement(26677. .26941)

/rpt_family="MIR"

repeat_region complement(27078. .27369)

/rpt_family="AluSx"

repeat_region complement(27374. .27682)

/rpt_family="AluSg"

repeat_region 27743. .28177

/rpt_family="LIM1/2"

repeat_region 28646. .28929

/rpt_family="AluSx"

repeat_region 28987. .29070

/rpt_family="(TA)n"

misc_feature 29085. .29241

REFERENCE	Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
AUTHORS	1 (bases 1 to 762)
TITLE	Schmitt, A., Specht, T., Dahl, E., Hinzmann, B., Rosenthal, A. and Pylarsky, C.
JOURNAL	Human nucleic acid sequences of bladder tumour tissue
	Patent: WO 99/54447-A 3 28-OCT-1999;
	SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
	BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
	(DE); PYLASKY CHRISTIAN (DE)
FEATURES	Location/Qualifiers
source	1..762
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
BASE COUNT	162 a 232 c 194 g 174 t
ORIGIN	
Alignment Scores:	
Pred. No.:	8 95e-125
Score:	117.00
Best Local Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	6 Gaps: 0
US-09-092-297-17 (1-117) x AX014140 (1-762)	
Oy	1 Proluenglproproarqalamealaproarglycysilevalalalaphelialle 20
Db	3 CCAGCTGCACACACCCAGACCAGCAGCTCCCGAGGCTGCATCTAGCTGTTCGCCATT 62
Oy	21 PheCysIISerArgrleuLeuCySSerHISglYAlaProValAlaProMetThrProTyR 40
Db	63 TTCTGATCTCCAGGCTCCTCTGCTCAGACGAGGCCCAAGTGCCCCCAATGATCCTTAC 122
Oy	41 LeuMetLeuCyGglProHISlysArqCysglYAspLysPheTyRAspProLeuGlnHIS 60
Db	123 CTGATGCTGTGCAGCCACACAGAAATGTGGGACAAAGTTCTACACACCCCTCGCAGCAC 182
Oy	61 CysCysTYrAspAspAlaValAlaValProLeuAlaArgThrGlnThrCysglYAsnCysThr 80
Db	183 TCTTCTAGTAAAGATGAGCCGCTGCTGCCCTTGCACAGACCCAGACGCTGTGGAAACTGCACC 242
Oy	81 PheArgrValCysPhegluGlnCysCysProTrpThrPheMetValLysLeuIleAsnGln 100
Db	243 TTCAGAGCTGCTTGGAGCAGATGCGCCCTCGACCTTCATGTAAGCTGATTAACAG 302
Oy	101 AsnCysAspSerAlaArgThrSerAspAspArgLeuCyArgSerValSer 117
Db	303 AACTCGACTCAGCCCGGACCTCGGATGACAGGCTTGTTCGACAGTGTACG 353
RESULT 2	
LOCUS	AX164131 771 bp DNA linear PAT 22-JUN-2001
DEFINITION	Sequence 3 from Patent WO0140465.
ACCESSION	AX164131
VERSION	AX164131.1 GI:14545080
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1 (bases 1 to 771)
AUTHORS	Fong, S., Goddard, A., Godowski, P. J., Grimaldi, C. J., Gurney, A. L.,
TITLE	Hillan, K. J., Tunas, D., Watanabe, C. K., Wood, W. L. and Zhang, Z.
JOURNAL	Compositions and methods for the treatment of immune related
	diseases
	Patent: WO 0140465-A 3 07-JUN-2001;
FEATURES	Genentech, Inc. (US)
source	1..771
	Location/Qualifiers
	/organism="Homo sapiens"
	/db_xref="taxon:9606"

BASE COUNT	169 a	231 c	195 g	176 t
ORIGIN				
Alignment Scores:				
Pred. No.:	9.05e-125	Length:	771	
Percent Similarity:	117.00%	Matches:	117	
Best Local Similarity:	100.00%	Conservative:	0	
Query Match:	100.00%	Mismatches:	0	
DB:	6	Indels:	0	
		Gaps:	0	
US-09-092-297-17 (1-117) x AXI64131 (1-771)				
QY	1	ProleuglnProrCraArgAlaMetaIAProArgLYcysIleValAlaVAlpheAalle	20	
Db	3	CCACTGCACACCACCGAGGCAATGGCTCCCCGAGGCTGCATCTACTGTCTTGGCAT	62	
QY	21	PhcCysIIeserArgIeuLeucyCysSerHISGLYAlaPProValAlaPromethrProtyr	40	
Db	63	TTCGTGATCTCCAGGCTCCTCTGCTCACACGAGGCCCAATGGCCCCCATGACTCTTCAC	122	
QY	41	LeuMetIeuCysGlInProHIsLysARgCYSGLYAsPLysPheTyRAsPProleuglInHS	60	
Db	123	CTGATCTCTGCGCAGGCACACAGAATGTGGGACAAGTTCTACACACCCCTTGCGACAC	182	
QY	61	CYSGYTYRASPSALaValAlaVAlProleualAargThrGIInThrcysglAnsCYstr	80.	
Db	183	TGTTGCTATGATATATCCCTGCGTGGCCCTTGCCAGACCCAGACGTGTGAATACTGCAC	242	
QY	81	PheArVAlCyaspheglInGlyncysCYSPROTPrphPhemetyAllysleulleansglIn	100	
Db	243	TTCAGAGCTGCTGTTTAGCAGATGCTGCCCTCGACCTTATGTGTGAAGCTGATAAACAG	302	
QY	101	ASnCYASPSeSaLaarGThrSerAsPaSParGleuCysArgSerValser	117	
Db	303	AACTGGAGCTCAGCCCGAGACCTCGATGACAGAGCTTTGTCGACATGTACAGC	353	
RESULT 3				
LOCUS	AC006262	107573 bp	DNA	linear PRI 21-SEP-2001
DEFINITION	Homo sapiens chromosome 19, BAC BC264576 (CTC-297N14), complete sequence.			
ACCESSION	AC006262			
VERSION	AC006262.2	GI:15718550		
KEYWORDS	HTG.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
	1 (bases 1 to 106320)			
	Lamerdin,J.E., McCreedy,P.M., Skowronski,E., Viswanathan,V.,			
	Burkhardt-Schultz,K., Gordon,L., Dias,J., Ramirez,M., Stillwagen,S.,			
	Phan,H., Velasco,N., Do,L., Regala,W., Terry,A., Gaines,J.,			
	Liu,S., Attis,C., Andreise,T., Trankhelm,M., Georgescu,A., Avila,J.,			
	Coerfeld,J., Duarte,S., Lucas,S., Bruce,R., Thomas,P., Quan,G.,			
	Kobayashi,I.B., Ariellano,A., Sanders,C., Ow,D., Nolan,M., Trong,S.,			
	Kobayashi,A., Olsen,A.S. and Carrano,A.V.			
	Sequence analysis of a 1.9 Mb region In 19q13.2 between APOE and			
	D19S412			
TITLE	Unpublished			
JOURNAL	2 (bases 1 to 107573)			
REFERENCE	Lamerdin,J.E.			
AUTHORS	Direct Submission			
TITLE	Submitted (30-DEC-1998) Joint Genome Institute, Lawrence Livermore			
JOURNAL	National Laboratory, 7000 East Ave., Livermore, CA 94551, USA			
REFERENCE	3 (bases 1 to 107573)			
AUTHORS	Lamerdin,J.E.			
TITLE	Direct Submission			
JOURNAL	Submitted (21-SEP-2001) Joint Genome Institute, Lawrence Livermore			
REFERENCE	National Laboratory, 7000 East Ave., Livermore, CA 94551, USA			
COMMENT	On Sep 21, 2001 this sequence version replaced gi:4079612			
	Map and sequence oriented from centromere to q telomere. BC264576			

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 3, 2002, 04:37:05 ; Search time 1825 Seconds

(without alignments)
1341.592 Million cell updates/sec

Title: US-09-092-297-17

Perfect score: 117
Sequence: 1 PUQPRAMPRGCIYAVPAI.....INQNDASRTSDRLCRYSV 117

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1797656 seqs, 10463268293 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3592932

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL-frame+pn.model -DEV-xlh
-Q/cgn2.1/USPTO.spool/US09092297/runat.29102002.094019.22469/app.query.fasta.1.263
-DB-GenEmbl -QFMT-fastap -SUFFIX-olip2n.rge -MINMATCH=0.1 -LOOCL=0 -LOOEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-oligo -TRANS-human40.cdi -LIST=45
-DOCALIGN=200 -THR.SCORER-quality -THR.MIN=1 -ALIGN=15 -MODE-LOCAL -OUTFMT-pto
-NORM-ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER-US09092297-QCEN.1.1.1182-erunat.29102002.094019.22469 -NCPV=6 -ICPV=3
-NO_XLPXY -NO_MAP -LARGSEQUERY -NEG.SCORER=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-VARN.TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :
GenEmbl.*
1: gb_da.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sy.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_da.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sy.*
28: em_un.*

29: em_vl.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117	100.0	762	6	AX014140 Sequence
2	117	100.0	771	6	AX164131 Sequence
3	82	70.1	107573	9	AC006262 Homo sapi
4	9	7.7	146180	9	AC007193 Homo sapi
5	9	7.7	158537	9	AC007193 Homo sapi
6	9	7.7	172246	9	HSAC002069 Human BAC
7	9	7.7	180563	2	AC024941 Homo sapi
8	9	7.7	228999	2	AC034170 Homo sapi
9	8	6.8	329	11	AC083885 Homo sapi
10	8	6.8	373	6	G00569 fruit fly S
11	8	6.8	792	1	AX309824 Sequence
12	8	6.8	2350	8	AF320338 Unculture
13	8	6.8	7637	9	AF369910 Arabidops
14	8	6.8	9366	2	HSCLCNC
15	8	6.8	12050	1	AC017452 Homo calcl
16	8	6.8	14178	2	AE002320 Neisseria
17	8	6.8	30107	2	AC014108 Drosophill
18	8	6.8	14932	1	AE009459 Brucella
19	8	6.8	42446	2	AC014407 Drosophill
20	8	6.8	55440	2	AF010283 Sorghum b
21	8	6.8	58856	8	AC020178 Drosophill
22	8	6.8	64885	2	AC004165 Arabidops
23	8	6.8	67297	3	AC110506 Mus muscu
24	8	6.8	74237	9	AC005135 Drosophill
25	8	6.8	74428	9	AL139283 Human DNA
26	8	6.8	86919	2	AL162495 Human DNA
27	8	6.8	96972	2	AC096809 Rattus no
28	8	6.8	103804	2	AC095725 Rattus no
29	8	6.8	105815	5	AP003472 Homo sapi
30	8	6.8	107344	5	AP002536 Oryza sat
31	8	6.8	123260	2	AC090119 Tefitugu
32	8	6.8	127508	2	AC109641 Homo sapi
33	8	6.8	128747	2	AP002964 Homo sapi
34	8	6.8	131353	9	AL060706 Oryza sat
35	8	6.8	132653	3	HS508115 Human DNA
36	8	6.8	133137	9	AC010015 Homo sapi
37	8	6.8	139330	9	AC007245 Homo sapi
38	8	6.8	140867	8	HS7824F16 Human DNA
39	8	6.8	151184	2	AP003216 Oryza sat
40	8	6.8	154894	9	AC098502 Rattus no
41	8	6.8	155213	9	AL035610 Homo sapi
42	8	6.8	155723	2	HSJ103J18 Human DNA
43	8	6.8	158895	3	AC009345 Oryza sat
44	8	6.8	162126	2	AC009345 Drosophill
45	8	6.8	162281	2	AC011784 Homo sapi
					AC011733 Homo sapi

ALIGNMENTS

RESULT 1
AX014140
LOCUS AX014140 762 bp
DEFINITION Sequence 3 from Patent WO9554447.
ACCESSION AX014140
VERSION AX014140.1 GI:10040587
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

DNA linear PAT 07-SEP-2000

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 3, 2002, 05:05:20 ; Search time 1686 Seconds

(without alignments)
936.621 Million cell updates/sec

Title: US-09-092-297-17

Perfect score: 117

Sequence: 1 PLOPRAMPKRCIVAVFAI.....INONCDANTSDDLRCRSVS 117

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 13736207 seqs, 674847542 residues

Word size: 1

Total number of hits satisfying chosen parameters: 27472246

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=x1h
-O=/cgn2.1/USPTO.spool/US09092297/runat.29102002.094020.22498/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=ol1p2n.rst -MINMATCH=0.1 -DOOPCL=0 -DOOPEXT=0
-INIT=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR.SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09092297@cgn.1.1.763 @runat.29102002.094020.22498 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6 -Fgapext=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	84	71.8	729 10 BG775059	BG775059 602650023
2	82	70.1	412 9 AA456370	AA456370 aa14e02.r

C	3	70	59.8	828	10	BG775668	BG775668 602650023
	4	58	49.6	201	10	BF911379	BF911379 IL2-UT007
	5	56	47.9	533	9	AM991605	AM991605 RCI-BN000
	6	24	20.5	275	9	A1572169	A1572169 t337e06.x
	7	16	13.7	475	9	AW795497	AW795497 MR4-UM001
	8	16	13.7	475	9	AA195677	AA195677 z332h04.s
	9	10	8.5	554	10	BE664072	BE664072 148225 MA
	10	9	7.7	403	10	F20995	F20995 HSPD05444 H
	11	9	7.7	527	9	BE035031	BE035031 MM01H04 M
	12	8	6.8	208	9	AW610810	AW610810 UP37905.x
	13	8	6.8	275	12	BH290179	BH290179 CH230-158
	14	8	6.8	340	12	AZ912795	AZ912795 RPCI-24-1
	15	8	6.8	415	10	BF912018	BF912018 IL2-UT007
	16	8	6.8	429	9	AM424984	AM424984 50784 MAR
	17	8	6.8	467	10	BE468930	BE468930 IPHAK0228
	18	8	6.8	531	12	AQ034057	AQ034057 I(3)03576
	19	8	6.8	551	9	BE131259	BE131259 I48-12647
	20	8	6.8	585	10	BM000341	BM000341 1031087C0
	21	8	6.8	599	12	AZ114409	AZ114409 RPCI-23-9
	22	8	6.8	600	10	BG966629	BG966629 602834475
	23	8	6.8	604	10	BG680336	BG680336 602629230
	24	8	6.8	609	10	BM301258	BM301258 MCR039D09
	25	8	6.8	609	12	BH268789	BH268789 CH230-21A
	26	8	6.8	609	12	BH357128	BH357128 CH230-106
	27	8	6.8	633	10	BJ073196	BJ073196 BJ073196
	28	8	6.8	661	12	AZ566931	AZ566931 228PVC05
	29	8	6.8	667	12	AO581989	AO581989 RPCI-11-4
	30	8	6.8	668	12	CNS02XNY	AL18599 Tetradon
	31	8	6.8	688	9	BB631694	BB631694 BB631694
	32	8	6.8	696	12	CNS01UE4	AL167701 Tetradon
	33	8	6.8	736	10	BG605861	BG605861 RH12.82-
	34	8	6.8	739	12	CNS01STL	AL165702 Tetradon
	35	8	6.8	771	10	BG357556	BG357556 60263878
	36	8	6.8	774	9	BE130713	BE130713 L48-91973
	37	8	6.8	800	12	BH269235	BH269235 CH230-56D
	38	8	6.8	805	9	BE034300	BE034300 MH02H09 M
	39	8	6.8	832	10	BE974238	BE974238 601680206
	40	8	6.8	868	12	CNS01TWB	AL173540 Tetradon
	41	8	6.8	887	9	BE035030	BE035030 MM01H03 M
	42	8	6.8	899	10	BE972075	BE972075 601651532
	43	8	6.8	904	12	CNS054S0	AL122115 Tetradon
	44	8	6.8	906	12	CNS02YMC	AL121937 Tetradon
	45	8	6.8	907	12	CNS04CF0	AL284373 Tetradon

ALIGNMENTS

RESULT 1
BG775059
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
TITLES
AUTHORS
JOURNAL
COMMENT

BG775059 729 bp mRNA linear EST 15-MAY-2001
602650023F1 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:4761029 5',
mRNA sequence.
BG775059
BG775059.1 GI:14045376
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 729)
NIH-MGC <http://mgc.ncl.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubln Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.llnl.gov>
Plate: L1CMI612 row: n column: 06

Location/Qualifiers
1. .729

BASE COUNT	169 a	212 c	181 g	166 t	1 others
ORIGIN					

BASE COUNT	88 a	116 c	119 g	89 t
ORIGIN				

US-09-092-297-17 (1-117) x AA456370 (1-412)

Db 275 GTCAGC 280

VERSION BG775668

REFERENCE 1 (bases 1 to 828)

ORGANISM Homo sapiens

[illegible]

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 533)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-rc1-BN0005-030500-015-907&t3=2000-03-03&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 16
High quality sequence stop: 533.
Location/Qualifiers
1..533
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BN0005"
/dev_stage="Adult"
/note="Organ: breast normal; Vector: puc18; Site:1: SmaI; Site:2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 116 a 156 c 150 g 111 t
ORIGIN
Alignment Scores:
Pred. No.: 4.8e-47 Length: 533
Score: 56.00 Matches: 56
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 47.86% Indels: 0
Gaps: 0
DB: 9
US-09-092-297-17 (1-117) x AW991605 (1-533)
QY 62 CysTYRASPAPALValValProLeuAlaArgThrGlnThrCysGlyAsnGlyThrPhe 81
Db 60 TGCATGATGATGCGCGTGTGCGCCCTGGCCAGACCCAGACGTGGGAACGACACCTTC 119
QY 82 ArgValCysPheGluGlnCysCysProThrPheMetValLysLeuIleasnGlnasn 101
Db 120 AGAGCTCTCTTGACCACTGCTGCCCTGGACCTTCATGCTAACTGATAAACAGAAC 179
QY 102 CysASPserAlaArgThrSerASPArgLysCysArgSerValSer 117
Db 180 TGGGCTCAGCCCGACCTCGATGACAGCGCTTGTCCAGTGTGAGC 227
RESULT 6
LOCUS A1572169 275 bp mRNA linear EST 13-APR-1999
DEFINITION t37e06.x1 Soares_NHMPU_s1 Homo sapiens cDNA clone IMAGE:2088898
3', mRNA sequence.

ACCESSION A1572169
VERSION A1572169.1 GI:4535543
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 275)
AUTHORS NCICGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 588 Std Error: 0.00
Seq primer: -40UP from G1600
High quality sequence stop: 272.
Location/Qualifiers
1..275
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:2088898"
/clone_lib="Soares_NHMPU_s1"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: p7773D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site:2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbHM, pregnant uterus NBHPU, and fetal heart NBH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."
BASE COUNT 58 a 77 c 56 g 84 t
ORIGIN
Alignment Scores:
Pred. No.: 8.16e-15 Length: 275
Score: 24.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.51% Indels: 0
Gaps: 0
DB: 9
US-09-092-297-17 (1-117) x A1572169 (1-275)
QY 41 LeuMetLeuGlnProHisLysArgCysGlyAspLysPheTYRASPProLeuGlnHis 60
Db 180 CTGATGCTGTGCACGACACAGAGATGTGGGACAAAGTTCTACGACCCCTCGACGAC 239
QY 61 CysCysTYRASP 64
Db 240 TGTTCATGAT 251
RESULT 7
LOCUS A1572169 130 bp mRNA linear EST 16-MAY-2000
DEFINITION MR4-UM0019-100400-203-e10 UM0019 Homo sapiens cDNA, mRNA sequence.
ACCESSION A1572169
VERSION A1572169.1 GI:7847367
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 130)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

TITLE Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brennan, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

JOURNAL MEDLINE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

COMMENT Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

CONTACT Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

TELEPHONE Tel: +55-11-2704922
Fax: +55-11-2707001

EMAIL Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-MR4-UM0019-100400-203-e106t3-2000-04-106t4-1)

SEQ PRIMER puc 18 forward
High quality sequence start: 19
High quality sequence stop: 129.

FEATURES
source
1..130
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="UM0019"
/dev_stage="Adult"
/note="Organ: uterus; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 23 a 53 c 25 g 29 t

ORIGIN

Alignment Scores:
Pred. No.: 4.82e-07 Length: 130
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.68% Indels: 0
DB: 9 Gaps: 0

US-09-092-297-17 (1-117) x AW795497 (1-130)

QY 33 ProValAlaProMetThrProTyrLeuMetLeuGlnProHisLys 48
|||||
Db 83 CCAGTGGCCCCGACCTGCTGATGATGCTGCGACGACACAG 130
|||||

RESULT 8
AA195677 475 bp mRNA linear EST 19-MAY-1997
LOCUS 2132104.S1 Soares_NhHMPU_S1 Homo sapiens cDNA clone IMAGE:665143
DEFINITION 3', mRNA sequence.
ACCESSION AA195677
VERSION AA195677.1 GI:11785355
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 475)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapelli, B., Chisose, S., Dietrich, N., Dubuque, T., Faveillo, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacey, M., Le, N., Mardis, E., Moore, B., Morris, K., Parsons, J., Prange, C., Riffkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisan, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Merra, M.

REFERENCE
AUTHORS

TITLE Generation and analysis of 280,000 human expressed sequence tags

JOURNAL MEDLINE Genome Res. 6 (9), 807-828 (1996)

COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

EMAIL Email: est@wustl.wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

INSERT LENGTH 857 Std Error: 0.00
High quality sequence stop: 308.

FEATURES
source
1..475
/organism="Homo sapiens"
/db_xref="GDB:5427175"
/db_xref="taxon:9606"
/clone_id="IMAGE:665143"
/clone_id="Soares_NhHMPU_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbHM, pregnant uterus NbHM, and fetal heart NbH19) were mixed, and as circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1 M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT 109 a 122 c 134 g 108 t 2 others

ORIGIN

Alignment Scores:
Pred. No.: 2.08e-06 Length: 475
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.68% Indels: 0
DB: 9 Gaps: 0

US-09-092-297-17 (1-117) x AA195677 (1-475)

QY 102 CysAspSerAlaArgThrSerAspArgLeuGlnArgSerValSer 117
|||||
Db 460 TGGGACTGACCGCGAGCTCGATGACAGCGCTTGTGCGAGTGCAC 413
|||||

RESULT 9
BE664072 554 bp mRNA linear EST 25-APR-2001
LOCUS 148225 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BE664072
ACCESSION BE664072
VERSION BE664072.1 GI:10023165
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
1 (bases 1 to 554)
Smith, T.P.L., Grose, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahnenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G., Perlea, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and Keefe, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)

REFERENCE
AUTHORS

COMMENT

Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smithemil1.marc.usda.gov
 Single pass sequencing. Bases called and alt-trimmed with phred
 v0.980904.e. Vector identified by cross-match with the -minscore 18
 and -fmaxmatch 12 options.

FEATURES

PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGTACAGCAG
 Plate: 51 row: M column: 8
 Seq primer: ATTGAGTGCACCTATAG.

SOURCE

Location/Qualifiers
 1. 554
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 4BOV"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: PCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from day 20 and day 40
 embryos."
 BASE COUNT 128 a 154 c 138 g 134 t
 ORIGIN

Alignment Scores:

Pred. No.: 3.14 Length: 554
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 8.55% Indels: 0
 Gaps: 0

US-09-092-297-17 (1-117) x BE664072 (1-554)

QY 76 CysgIyAaNCysThrPheArgValCysPhe 85
 ||||||||||||||||||||||||||||||||
 Db 222 TGTGGAGACTGCACCTTATGAGGTCTGCTTC 251

RESULT 10

F20995 403 bp mRNA linear EST 17-MAY-1999
 LOCUS HSPD05444 HM3 Homo sapiens cDNA clone 034-X4-13, mRNA sequence.
 DEFINITION F20995
 ACCESSION F20995.1 GI:2060171

VERSION F20995.1 GI:2060171
 KEYWORDS EST.

ORGANISM

Human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 403)
 Lanfranchi, G., Muraro, T., Caldara, F., Pacchioni, B., Pallavicini, A.,
 Pandolfo, D., Toppo, S., Trevisan, S., Scarso, S. and Valle, G.

IDENTIFICATION OF 4370 EXPRESSED SEQUENCE TAGS FROM A
 3'-END-SPECIFIC cDNA LIBRARY OF HUMAN SKELETAL MUSCLE BY DNA
 SEQUENCING AND FILTER HYBRIDIZATION

JOURNAL MEDLINE
 Genome Res. 6 (1), 35-42 (1996)
 96276048

COMMENT
 Contact: Valle G.
 CRI Biotechnology Centre
 University of Padua
 Via Trieste 75, 35121 Padua, Italy
 ABI Chromatograms and other information are available on WWW at
 http://grup.bio.unipd.it.

FEATURES
 Location/Qualifiers
 1. 403

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="X4-X4-13"
 /clone_id="HM3"
 /sex="female"

/tissue_type="pectoral muscle (after mastectomy)"

/note="Vector: pcDNA1 (Invitrogen); Site_1: BstXI;
 Site_2: NotI. This library was constructed by G.

Lanfranchi. This library is not subtracted nor normalized.
 The first strand cDNA was primed with a biotinylated
 oligo-dT-NotI primer

(5'-Biotin-AACCGGCTCAGAGCGCCGCTTTTCTTTTCTTTT-3'). The
 ds cDNA was sonicated and size-selected in the range
 350-550 bp. The 3' specific fragments were selected by
 streptavidin coated magnetic beads, ligated to
 non-palindromic BstXI adapters, NotI digested and
 directionally cloned into BstXI-NotI cut pcDNA1 vector."

BASE COUNT 105 a 97 c 98 g 97 t 6 others
 ORIGIN

Alignment Scores:

Pred. No.: 22.8 Length: 403
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.69% Indels: 0
 Gaps: 0

US-09-092-297-17 (1-117) x F20995 (1-403)

QY 66 AlaValProLeuAlaArgThrGln 74
 ||||||||||||||||||||||||||||
 Db 297 GCCGTGTGCTCTCGCAGACACAA 271

RESULT 11

BE035031 527 bp mRNA linear EST 07-JUN-2000
 LOCUS M001H04 MM Mesembryanthemum crystallinum cDNA 5' similar to
 DEFINITION BE035031
 elongation factor 1-beta, mRNA sequence.
 ACCESSION BE035031.1 GI:8330040

VERSION BE035031.1 GI:8330040
 KEYWORDS EST.

ORGANISM

Common ice plant.
 Mesembryanthemum crystallinum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Caryophyllidae; Caryophyllales; Alzaceae; Mesembryanthemum.

REFERENCE 1 (bases 1 to 527)
 Bonner, H.J., Borchert, C., Brazille, S., Brooks, J., Eaton, M., Ferreira,
 H., Kawasaki, S., McCollough, A., Michalowski, C.B., Palacio, C.,
 Scara, G., Wheeler, M. and Zepeda, G.R.

FUNCTIONAL GENOMICS OF PLANT STRESS TOLERANCE
 Unpublished (2000)
 Contact: Michalowski, C.B.

JOURNAL COMMENT
 University of Arizona
 Bio Sciences West room 513, Tucson, AZ 85721, USA
 Tel: 520-621-7982
 Fax: 520-621-1697
 Email: chm@u.arizona.edu

FEATURES
 Location/Qualifiers
 1. 527

/organism="Mesembryanthemum crystallinum"
 /db_xref="taxon:3544"
 /clone_lib="NM"
 /cell_type="epidermal bladder cells"
 /dev_stage="12 weeks old"
 /note="Vector: Bluescript SK+; Site_1: EcoRI; Site_2: XhoI
 ; Plants stressed 6 weeks in 500mM NaCl"

BASE COUNT 148 a 111 c 149 g 119 t
 ORIGIN

Alignment Scores:

Pred. No.: 30.9 Length: 527
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.69% Indels: 0

AUTHORS
Zhao, S., Niernan, W., Malek, J., Shartsman, S., Aklnret, B., Levins, M., Tsegaye, G., Geer, K., Krol, M., Shartsbeyn, A., Gebregeorgis, E., Russell, D., de Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Other_GSSs: RPCI-24-163p16.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

TITLE
JOURNAL
COMMENT

FEATURES
source
Location/Qualifiers
1. 340
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-163p16"
/clone_lib="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTRABAC1; Site 1: BamHI; Site 2: BamHI; RPCI-24 Mouse BAC library produced by Pieter de Jong. The library was cloned in the pTRABAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."
BASE COUNT 94 a 79 c 81 g 86 t
ORIGIN

Alignment Scores:
Pred. No.: 196 Length: 340
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.84% Indels: 0
DB: 12 Gaps: 0

US-09-092-297-17 (1-117) x AC912795 (1-340)

QY 68 ValProLeuAlaArgThrGlnThr 75
|||||
Db 131 GTTCCATTGCTCGCACACACACA 154

RESULT 15.
BF912018 415 bp mRNA linear EST 18-JAN-2001
LOCUS IL2-UTR0073-121100-232-D01 UTR0073 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF912018
ACCESSION BF912018
VERSION BF912018.1 GI:12303476
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 415)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Baia, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=IL2&t=IL2-UTR0073-121100-232-D01&t3=2000-11-12&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 409.

FEATURES
source
Location/Qualifiers
1. 415
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UTR0073"
/dev_stage="Adult"
/note="Organ: uterus; tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from OREGRES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 89 a 78 c 111 g 137 t
ORIGIN

Alignment Scores:
Pred. No.: 245 Length: 415
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.84% Indels: 0
DB: 10 Gaps: 0

US-09-092-297-17 (1-117) x BF912018 (1-415)

QY 20 IlePheCysIleSerArgLeu 27
|||||
Db 354 ATATCTGCAATTCGCCGTTACTTA 377

Search completed: November 3, 2002, 07:00:18
Job time: 1693 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 3, 2002, 03:06:09 ; Search time 1686 Seconds
(without alignments)
936.621 Million cell updates/sec

Title: US-09-092-297-17

Perfect score: 668
Sequence: 1 PLOPPRAMPRCIVAVFAI.....INONCSARTSDRLCRSVS 117

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapext 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ .p2n.model -DEV-xml
-Q/cg2_1/USPTO.spool/US09092297/rnat_29102002_093923_21777/app_query.fasta_1.263
-DB-est -QPMF-fastap -SUFFIX-p2n.rst -MINMATCH-0.1 -LOOPEXT-0
-UNITS-bits -START-1 -END-1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST-45
-DOCALLGN-200 -THR-SCORE-pct -THR-MAX-100 -THR-MIN-0 -ALIGN-15 -MODE-LOCAL
-OUTFMT-pco -NORF-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-200000000
-USER-US09092297_@CGN_1.1_763_@rnat_29102002_093923_21777 -NCPV-6 -ICPU-3
-NO_XLPEX -NO_MMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -LONGLOG -DEV-TIMEOUT-120
-NARN-TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6 -FGAPEXT-7
-YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database :

EST:.*
1: em_estba:.*
2: em_esthum:.*
3: em_estnu:.*
4: em_estnu:.*
5: em_estov:.*
6: em_estpl:.*
7: em_estro:.*
8: em_hic:.*
9: gb_est1:.*
10: gb_est2:.*
11: gb_hic:.*
12: gb_gss:.*
13: em_gss_hum:.*
14: em_gss_huv:.*
15: em_gss_pln:.*
16: em_gss_vrt:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	553	82.8	729	10	BG775059
2	536	80.2	828	10	BG775668

3	513	76.8	412	9	AA456370
4	408.5	61.2	554	10	BE664072
5	389.5	58.3	555	10	BE664075
6	358.5	53.7	533	9	AW991605
7	340	50.9	201	10	BF911379
8	236.5	35.4	366	9	BE122850
9	210.5	31.5	537	9	AA024389
10	190.5	28.5	481	9	AA1806131
11	183.5	27.5	523	9	AA456254
12	180.5	27.0	275	9	AT572169
13	178.5	26.7	455	9	AA454642
14	177.5	26.6	464	9	AT275379
15	170.5	25.5	454	9	AA1081371
16	150	22.5	575	9	AA760247
17	124.5	18.6	422	9	AT263445
18	102	15.3	475	9	AA195677
19	97	14.5	396	10	BJ124893
20	97	14.5	578	10	BJ104084
21	97	14.5	580	10	BJ115213
22	97	14.5	637	10	BJ103226
23	97	14.5	640	10	BJ112066
24	97	14.5	650	10	BJ121757
25	94	14.1	1288	10	BE906149
26	93.5	14.0	609	12	AZ104501
27	93	13.9	130	9	AW795497
28	92	13.8	482	10	BJ058488
29	92	13.8	713	12	AQ256383
30	92	13.8	1461	10	BF341539
31	90.5	13.5	953	10	BF138520
32	90	13.5	360	9	AV193603
33	90	13.5	1176	12	AG115743
34	89.5	13.4	601	10	BJ127189
35	89.5	13.4	1156	12	AG160702
36	89.5	13.4	1157	10	BG244032
37	89.5	13.4	1168	10	BG387111
38	89	13.3	567	9	AA122680
39	89	13.3	594	12	AZ219997
40	89	13.3	622	12	AZ361346
41	88.5	13.2	1575	10	BF970881
42	87.5	13.1	443	9	AT853580
43	87.5	13.1	523	9	AM556673
44	87.5	13.1	658	10	BG075243
45	87.5	13.1	878	10	BI949596

ALIGNMENTS

RESULT 1
BG775059
LOCUS 729 bp mRNA linear EST 15-MAY-2001
DEFINITION 602650023F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4761029 5',
mRNA sequence.
ACCESSION BG775059
VERSION BG775059.1 GI:14045376
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 729)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM1612 row: n column: 06

High quality sequence stop: 726.

FEATURES	Location/Qualifiers
source	1. .729

BASE COUNT	169 a	212 c	181 g	166 t	1 others
ORIGIN					

Alignment Scores:	
Pred. NO.:	4.09e-47
Score:	553.00
Percent Similarity:	98.99%
Best Local Similarity:	97.98%
Query Match:	82.78%
DB:	10
	Gaps:
	0

QY 40 TyrIleuMetIeuCysGlnProHlaLysArgCysGlyLysPlySerPheTyrAspProLeuGln 59
 Db 61 TACTGTATGCTGCTCCAGCCACACAAAGATGTGGGGAAGAAGTCTACGACCCCCCTCAG 120
 QY 60 HisCysCysTyrAspAspAlaValaValaProLeuAlaIarGthr-GlnThrCysGlyAsnCy 79
 Db 121 CACTGTGCTATGATGATGATGCCGTGTGCTCCCTTGGCCAGGACCCACAGACGTGTGMAATCG 180
 QY 79 sThrPheArgValCysPheGluGlnCysCysProTyrPheMetValLysLeuIleAs 99
 Db 181 CACCTTAGAGTCTGCTTATAGCAGTCTCCCTCGACCTTATGTTGTAAGCTGATATAA 240
 QY 99 ngInAsnCysAspSerAlaIarGthrSerAspAspArgLeuCysArgSerAlaSer 117
 Db 241 CCAGAACTGGCAGTACAGCCCGGAGACCTTGATGACAGGCTTTGTTCGCAGTCTCAGC 295

RESULT 2	Bg775668/c	828 bp	mRNA	linear	EST 15-MAY-2001
LOCUS	Bg775668/c				
DEFINITION	60265002T1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4761029 3', mRNA sequence.				
ACCESSION	Bg775668				
VERSION	Bg775668.1	GI:14045985			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	1 (Bases 1 to 828)				
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				

found through the I.M.A.G.E. Consortium/LLNL at
http://image.llnl.gov
Plate: LLCM1612 row: n column: 06
High quality sequence start: 16
High quality sequence stop: 806.

FEATURES	Location/Qualifiers
source	1. . 828

BASE COUNT	190 a	214 c	234 g	190 t
ORIGIN				

```

Alignment Scores:
Pred. No.:      2.74e-45      Length:      828
Score:          536.00      Matches:      96
Percent Similarity: 98.98%      Conservative: 1
Best Local Similarity: 97.96%      Mismatches:  1
Query Match:    80.24%      Indels:      0
DB:             10      Gaps:        0
US-09-092-297-17 (1-117) x BG775668 (1-828)

```

QY 40 TyrLeuMetIleuCyGlnProHisLysArgCysGlyAspLysPheTyrAspProLeuGln 59

Db 559 TACCAGTAGCTGTGCACACCACACAGAGATGTGGGGCAAACTTTACACACCCCTTCG 540

QY 60 HisCysCysTyrAspAspAlaValAlaProLeuAlaArgHisGlnThrCysGlyAsnCys 79

Db 539 CACTGTTCATGATGATGATGCCGTGTGCCCTTGGCCAGGACCCACAGCTGTGAATGCC 480

QY 80 ThrPheArgValCysPheGlnGlnCysCysProTrpThrPheMetValLysLeuIleAsn 99

Db 479 ACCCTCAAGAGCTGCTTTGACACAGTCGCCCC-TGGACCTTCATGTGTCAGCTGATAAAC 421

QY 100 GlnAsnCysAspSerAlaArgThrSerAspAspArgLeuCysArgSerValSer 117

Db 420 CAGAACTCGACTACGCCCGGAGCCTCGGATGACAGGCTTTTGCAGAGTCTACG 367

AA456370	AA456370	412 bp	mRNA	linear	EST 06-JUN-1997			
LOCUS	aa44602.t1	Soares_NhmPu_Si	Homo sapiens	CDNA clone	IMAGE:813242			
DEFINITION	5' mRNA sequence.							
ACCESSION	AA456370							
VERSION	AA456370.1	GI:2178946						
KEYWORDS	EST.							
SOURCE	human.							
ORGANISM	Homo sapiens							

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (pages 1 to 412)	Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Joist, S., Knacab, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, R., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, A., Waterston, R. and Wilson, R.	WashU-Merck EST Project 1997	Unpublished (1997)

COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (infoimage.lnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 398.

FEATURES

source

1. 412
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDB:6044152"
/db_xref="taxon:9606"
/clone_lib="IMAGE:813242"
/clone_lib="Soares_NhHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NDM, pregnant uterus
NBHPU, and fetal heart NBH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

BASE COUNT

88 a 116 c 119 g 89 t

ORIGIN

Alignment Scores:

Pred. No.: 2,49e-43 Length: 412
Score: 513.00 Matches: 92
Percent Similarity: 98.928 Conservative: 0
Best Local Similarity: 98.928 Mismatches: 1
Query Match: 76.804 Indels: 1
DB: 9 Gaps: 0

US-09-092-297-17 (1-117) x AA456370 (1-412)

QY 25 AAGLEuLeucysSerHisglYAlaProValAlaPrometHrProTyrluMetLeucys 44
|||
DB 3 AGGCTCTCTGCGACACGAGGCCCGACAGTGGC-CCCATGACTCTTACCTGATGCTGTC 61
|||
QY 45 GlnProHisLysArgCysglYAspLysPheTyAspProLeuGlnHisCysCysTyrlAsp 64
|||
DB 62 CAGCCACACAGAGATGTGGGACAGATCTTACGACCCCTGACAGCACTGTCTATGAT 121
|||
QY 65 AspaLaValAlaProLeuAlaArgThrGlnThrCysglYAsnCysThrPheArgValCys 84
|||
DB 122 GATGCCGTGTCGCCCTGGCCAGGACCCAGAGCTGTGGAACCTGCACCTTCAAGTCTGC 181
|||
QY 85 PhegluGlnCysCysProTrpThrPheMetValLysLeuLLeasnGlnAsnCysAspSer 104
|||
DB 182 TTGAGCACTGTGTCGCCCGGACCTTCATGCGAGAGCTATTAACCAAGACTCGACTCA 241
|||
QY 105 AlaArgThrSerAspaPArgLysCysArgSerValSer 117
|||
DB 242 GCCCGGACCTCGGATGACAGGCTTTGTCTGCAGTGTACG 280
|||
RESULT 4 BE664072 554 bp mRNA linear EST 25-APR-2001
LOCUS BE664072
DEFINITION 148225 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE664072
VERSION BE664072.1 GI:10023165
KEYWORDS EST.
SOURCE COV.
ORGANISM Bos taurus

REFERENCE

1 (bases 1 to 554)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,
G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
Perte,G., Holt,I., Karaycheva,S., Liang,F., Quackenbush,J. and
Keefe,J.W.

TITLE

Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle

JOURNAL

Genome Res. 11 (4), 626-630 (2001)

MEDLINE

21180013

COMMENT

Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smithemall.marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross-match with the -minscore 18
and -mismatch 12 options.
PCR primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCGCCATCGACGACG
Plate: 51 row: M column: 8
Seq primer: ATTGAGTGCATATAG.
Location/Qualifiers
1. 554
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 4BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORF6; Site_1: XbaI, Site_2: XhoI;
Library made from pooled tissue from day 20 and day 40
embryos."

FEATURES

source

BASE COUNT 128 a 154 c 138 g 134 t
ORIGIN

Alignment Scores:

Pred. No.: 2,19e-32 Length: 554
Score: 408.50 Matches: 72
Percent Similarity: 73.918 Conservative: 13
Best Local Similarity: 62.618 Mismatches: 23
Query Match: 61.158 Indels: 7
DB: 10 Gaps: 1

US-09-092-297-17 (1-117) x BE664072 (1-554)

QY 6 ArgAlaMetAlaProArgLysIleValAlaValPheAlaIlePheCysIleSerArg 25
|||
DB 12 AGAGCATGATCTCCACGACCTGCTGCTGTGGCCGCTGTGTCATCTCTTTGG 71
|||
QY 26 LeuLeucysSerHisglYAlaProValAlaPrometHrProTyrluMetLeucysGln 45
|||
DB 72 CTCTCTGTTCACATGCTGTGCCAGTGTCCCGACAGGACCTAGCTGTTCCTGCCAG 131
|||
QY 46 ProHisLysArgCysglYAspLysPheTyAspProLeuGlnHisCysCysTyrlAsp 65
|||
DB 132 TCACAGAGGATGCGGGGACCGATTCTACGACCCCGGACAGACTGTGTGATGACAT 191
|||
QY 66 AlaValAlaProLeuAlaArgThrGlnThrCysglYAsnCysThrPheArgValCysPhe 85
|||
DB 192 GCCATGTACCTCTGGGCGACGACCCGGAAGTGTGGAGCTGCACCTTTAGGCTGTCTTC 251
|||
QY 86 GluGlnCysCysProTrp-----ThrPheMetValLysLeuIle 98
|||
DB 252 AAGCAGTGTCTGCGCGGTGGTGCACAGGCCCGCAGAGACTCTCTGCTGTGAAGTGAA 311
|||
QY 99 AsnGlnAsnCysAspSerAlaArgThrSerAspArgValCys 113
|||
DB 312 GGTCAAAATTTTACTCGGGTGTATCTCAGACGATCGGGTTTGT 356
|||

```

RESULT 5
LOCUS BE664075 555 bp mRNA linear EST 25-APR-2001
DEFINITION 148331 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE664075
VERSION BE664075.1 GI:10023171
KEYWORDS EST.
SOURCE
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE
AUTHORS 1 (bases 1 to 555)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G.,
Peters,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keefe,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
JOURNAL MEDLINE 21180013
COMMENT
Contact: Smith TPL
USDA, ARS US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross-match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTCCAGTACGACG
Plate: 51 row: N column: 8
Seq primer: ATTAGTGACACTATG.
FEATURES
source
location/Qualifiers
1..555
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 4BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
library made from pooled tissue from day 20 and day 40
embryos."
BASE COUNT 130 a 157 c 133 g 135 t
ORIGIN
Alignment Scores:
Pred. No.: 1,99e-30 Length: 555
Score: 389.50 Matches: 70
Percent Similarity: 71.30% Conservative: 12
Best Local Similarity: 60.87% Mismatches: 26
Query Match: 58.31% Indels: 7
DB: 10 Gaps: 1
US-09-092-297-17 (1-117) x BE664075 (1-555)
QY 6 ArgAlaMetAlaProArgGlyIleValAlaValAlaPheAlaIlePheCysIleSerArg 25
Db 12 AAGAGCCATAGTCACAGAGCGATCCGCTGACTGCGCTGTCTGATCCCTTTG 71
QY 26 LeuLeuCySerHisGlyValaProValaIaProMetThrProTyrLeuMetLeuCysGln 45
Db 72 CTCTCTGTTCACATGCTGCCACAGTCTCCACACAGGACATCAGCTGTGCTGTGCCAG 131
QY 46 ProHisLysArgCysGlyAspLysPheTyrAspProLeuGlnHisCysCysTyrAspAsp 65
Db 132 TCACAAGACGATGCGGGGAGCATTTCTACGACCCCGGACAGAGACTGTTCATATGACAT 191
QY 66 AlaValaIaProLeuAlaIaArgThrGlnThrCysGlyAsnCysTyrPheArgValCysPhe 85

```

```

Db 192 GCCATGTACCTTGGGAGAGACCCGGAAGTGTGGAACTGCACCTTACGCTCTCTC 251
QY 86 GluGlnCysCysProTyrP-----ThPheMetValysLeuIle 98
Db 252 GAGACATGCTGCCCGGTGGTTGGTCAACAGCCCCCATGATCTCTTCTGTGTGAAGTGAAA 311
QY 99 AsnGlnAsnCysAspSerAlaArgThrSerAspAspArgLeuCys 113
Db 312 GGTCAAAATTTGTTACTCGCGTGTATCTCATACAGATGGGTTTCT 356
RESULT 6
LOCUS AW991605 533 bp mRNA linear EST 05-JUN-2000
DEFINITION RCI-BN0005-030500-015-g07 BN0005 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW991605
VERSION AW991605.1 GI:8251660
KEYWORDS EST.
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 533)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Brlones,M.R.,
Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.G.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL MEDLINE 20202663
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=et2-RCI-BN0005-030
500-015-g07&ls=2000-05-03&ls=1)
Seq primer: puc 18 forward
High quality sequence start: 16
High quality sequence stop: 533.
FEATURES
source
location/Qualifiers
1..533
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BN0005"
/dev_stage="Adult"
/note="Organ: breast_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 116 a 156 c 150 g 111 t
ORIGIN
Alignment Scores:
Pred. No.: 2,95e-27 Length: 533
Score: 358.50 Matches: 69
Percent Similarity: 94.52% Conservative: 0
Best Local Similarity: 94.52% Mismatches: 4
Query Match: 53.67% Indels: 3
DB: 9 Gaps: 0
US-09-092-297-17 (1-117) x AW991605 (1-533)

```

OY 45 GlnProHISLysArgCysGlyAspLysPheTyrAspProLeuGlnHIScysCysTyrAsp 64
 |||||||
 DB 12 CAGCCACACAGAGATGTGGGAGACG-TTCTACGGCCCCCT-CAGCAGTGG-TGCTATGAT 68
 OY 65 AspAlaValValProLeuAlaArgThrGlnHIScysGlyAspCysThrPheArgValcys 84
 |||||||
 DB 69 GATGCCCTGTGGCCCTTGGCCAGAGACCAGCTGTGAAACTGGACCTTCAGACTCTGC 128
 OY 85 PheGluGlnCysCysProTyrThrPheMetValLysLeuIleAsnGlnAsnCysAspSer 104
 |||||||
 DB 129 TTGAGCAGTCTGCTGCCCCCTGACCTTCATGTGGAGCTGATTAACAGACATCGACATCA 188
 OY 105 AlaArgThrSerAspAspArgLeuLysArgSerValser 117
 |||||||
 DB 189 GCCCGACCTCGATGACAGAGCTTTGTCGACGTTCAGC 227
 RESULT 7
 LOCUS BE911379 201 bp mRNA linear EST 18-JAN-2001
 DEFINITION U12-UT0073-081100-213-D02 UT0073 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE911379
 VERSION BE911379.1 GI:12302837
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 201)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Bionesi,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 JOURNAL 2020663
 MEDLINE Contact: Simpson A.J.G.
 COMMENT Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=U12&t2=U12-UT0073-
 081100-213-D02&t3=2000-11-08&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 13
 High quality sequence stop: 201.
 Location/Qualifiers
 1..201
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="UT0073"
 /dev_stage="Adult"
 /note="Organ: uterus-tumor; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ONESTEP PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 39 a 63 c 51 g 48 t
 ORIGIN
 Alignment Scores: 6,33e-26 Length: 201
 Pred. No.: 340.00 Matches: 58
 Score:

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 50.90% Indels: 0
 DB: 10 Gaps: 0
 US-09-092-297-17 (1-117) x BE911379 (1-201)
 OY 55 TyrAspProLeuGlnHIScysCysTyrAspAspAlaValProLeuAlaArgThrGln 74
 |||||||
 DB 28 TACGACCCCTGCACACACTGTTGCTATGATGATGCCGTCGTCCCTTGGCCAGAGCCAG 87
 OY 75 ThrCysGlyAsnCysThrPheArgValcysPheGluGlnCysCysProTyrThrPheMet 94
 |||||||
 DB 88 ACGTGTGAACCTGCACCTTCAGACTCTGTTGAGCAGTCTGCCCCCTGACCTTCATG 147
 OY 95 ValLysLeuIleAsnGlnAsnCysAspSerAlaArgThrSerAspAspArgLeu 112
 |||||||
 DB 148 GTAAAGCTGATTAACAGAACTGCAGTCAGCCGACCTGAGATGACAGCTT 201
 RESULT 8
 LOCUS BE122850/c 366 bp mRNA linear EST 26-OCT-2000
 DEFINITION 02.15 Human Epidermal Keratinocyte Subtraction Library- Upregulated
 Transcripts Homo sapiens cDNA clone 02.15 5' similar to Homo
 sapiens chromosome 19, mRNA sequence.
 ACCESSION BE122850
 VERSION BE122850.1 GI:8514955
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 366)
 Avery,D.P. and Clark,D.H.
 Human Epidermal Keratinocyte Subtraction Library- Upregulated
 Transcripts
 Unpublished (2000)
 JOURNAL Contact: Schlager,J.J.
 COMMENT Molecular Toxicology
 United States Army Medical Research Institute of Chemical Defense
 (USAMRICD)
 3100 Ricketts Point Road; ATTN: MCMR-DV-PA; APG-EA, MD 21010-5400,
 USA
 Tel: 410 436 1940
 Fax: 410 436 1960
 Email: John.Schlager@AMEDD.ARMY.MIL
 Seq primer: Forward Seq Primer: GCCAGTGTGATGATATCTG; Reverse Seq
 Primer: GCTCGATCCACATGTAACG
 High quality sequence stop: 366
 POLY(A)-NO.
 Location/Qualifiers
 1..366
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="02.15"
 /clone_lib="Human Epidermal Keratinocyte Subtraction
 Library- Upregulated Transcripts"
 /sex="Female"
 /tissue_type="skin"
 /cell_type="Primary keratinocyte"
 /cell_line="Clonetics 4075"
 /dev_stage="adult"
 /lab_host="E. coli (ampicillin-resistant)"
 /note="Vector: pT-Adv; Library preparation: Sequence
 isolated using PCR-select (Clontech) subtraction library
 construction (Datchenko et al., (1996) Proc. Natl. Acad.
 Sci. USA 93: 6025-6030) after sulfur mustard exposure of
 primary human adult epidermal keratinocytes. Subtraction
 library cloned for random sequence selection into pT-Adv
 vector."
 BASE COUNT 101 a 92 c 103 g 70 t
 ORIGIN

/organism="Homo sapiens"

/db.xref="taxon:9606"
/clone_1lb="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site: 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBH19W, testis NHT, and B-cell MCL-GAP_GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 120 a 108 g 135 t
ORIGIN

Alignment Scores:
Pred. No.: 5,28e-10 Length: 481
Score: 190.50 Matches: 41
Percent Similarity: 58.14% Conservative: 9
Best Local Similarity: 47.67% Mismatches: 23
Query Match: 28.52% Indels: 13
DB: 9 Gaps: 4

US-09-092-297-17 (1-117) x A1806131 (1-481)

43 LeucylGlnProHisLysArgCysGlyAspLysPheTyrAspProLeuGlnHisCysCys 62
|||||
465 CTGTGGCAGCCGACACCCAGGTGTGGAGACAAAGATCTACAAACCCCTGGAGCACTGCTGT 406
63 TyrAspAlaValAlaProLeuAlaArgTglnThrCysGly---AsnCysThrPhe 81
|||||
405 TACAATGAGCCCATCGTGTCCCGAGGAGAACCCGCAATGTGTCTCCCTGCACCTTC 346
82 ArgValCysPheGlnGlnCysCys-----ProTyrPheMetVal 95
|||||
345 TGGCCCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 286
96 LysLeu-----IleAsnGlnAsnCysAspSerIleArgTyrIleSerAspArg 111
|||||
285 AACCTAAGAGTTCAGGCTGTGATGTCACACTGCTCCATCTCCATCTCCAGTAAG--- 229
112 LeucylArgSerValSer 117
|||
228 ---TGTGAAGCAGCAACT 214

RESULT 11
AA456254 523 bp mRNA linear EST 06-JUN-1997
LOCUS zx99d06.r1 Soares.NHMPu_S1 Homo sapiens cDNA clone IMAGE:811883
DEFINITION 5', mRNA sequence.
ACCESSION AA456254
VERSION AA456254.1 GI:2179464
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 523)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Getse, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,
Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wyllie,
T., Waterston, R. and Wilson, R.
WashU-Werck EST Project 1997
Unpublished (1997)
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (lnl@image.lnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amsterdam
High quality sequence stop: 497.
Location/Qualifiers
1. 523

FEATURES
Source
/organism="Homo sapiens"
/db.xref="GDB:6042794"
/db.xref="taxon:9606"
/clone_1lb="Soares_NHMPu_S1"
/clone_1lb="Soares_NHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site: 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NDHM, pregnant uterus NBHPU, and fetal heart NBH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT 99 a 152 c 140 g 132 t
ORIGIN

Alignment Scores:
Pred. No.: 3.11e-09 Length: 523
Score: 183.50 Matches: 45
Percent Similarity: 52.73% Conservative: 13
Best Local Similarity: 40.91% Mismatches: 39
Query Match: 27.47% Indels: 14
DB: 9 Gaps: 5

US-09-092-297-17 (1-117) x AA456254 (1-523)

10 ProArgLysCysIleValAlaValPheAlaIlePheCysIleSerArgLeuLysSer 29
|||||
200 CCCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 256
30 HisGlyAlaProValAlaProMetThrProTyrLeuMetLeuGlnProHisLysArg 49
|||||
257 AGGGAAGTC---ATCGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312
50 CysGlyAspLysPheTyrAspProLeuGlnHisCysCysTyrAspAlaValAlaPro 69
|||||
313 TGTGGAGACAAGATCTACAAACCCCTGGAGACAGTGTGTTACATGAGCCGATGCTGTC 372
70 LeuAlaArgTglnThrCysGly---AsnCysThrPheArgValCysPheGlnGlnCys 88
|||||
373 CTGAGCGAGACCCGCAATGTGTGCTCCCGCTGACCTTCTGCGCTTGTGAGCTGTCG 432
89 Cys-----ProTyrPheMetValLysLeu-----Ile 98
|||||
433 TGTCTGATTCCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 492
99 AsnGlnAsnCysAspSerAlaArgThrSer 108
|||
493 AATTCACAGTCCACATCTCCATCTCC 522

RESULT 12
A1572169 275 bp mRNA linear EST 13-APR-1999
LOCUS test7e06.x1 Soares.NHMPu_S1 Homo sapiens cDNA clone IMAGE:2088888
DEFINITION 3', mRNA sequence.
ACCESSION A1572169
VERSION A1572169.1 GI:4535543
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 464)
TITLE NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapb-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 355 Std Error: 0.00
Seq primer: -400P from Gldco
High quality sequence stop: 436.
Location/Qualifiers
1. 464
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1876980"
/issue_type="Soares_NHMPu.S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: p773D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2Nbhm, pregnant uterus
Nbhpv, and fetal heart Nbhm19v) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479." 1 others

BASE COUNT 115 a 113 c 108 g 127 t

ORIGIN

Alignment Scores:
Pred. No.: 1,1e-08 Length: 464
Score: 177.50 Matches: 38
Percent Similarity: 55.42% Conservative: 8
Best Local Similarity: 45.78% Mismatches: 24
Query Match: 26.57% Indels: 13
DB: Gaps: 4

US-09-092-297-17 (1-117) x A1275379 (1-464)

QY 44 CysGlnProHisLysArgCysGlyAspLysPheTyrAspProLeuGlnHisCysCysTyr 63
|||||
Db 462 TGGCAGCGGCGACACAGAGTGTGAGACAGATCTCACCCCTGGAGAGCATGTGTTAC 403
|||||

QY 64 AspAspAlaValValProLeuAlaArgThrGlnThrCysGly---AsnCysThrPheArg 82
:::|||||
Db 402 AATGACGCCCATCTGCTCCGAGGAGACCCGCCAATGTGTTGCCCTGCACCTCTGG 343
|||||

QY 83 ValCysPheGlnGlnCysCys-----ProTyrPheMetValLys 96
|||||
Db 342 CCTGTGCTTACGCTGCTCTTGTGATCTTGGCTTCACAAAGATTTGTTGTTGAAG 283
|||||

QY 97 Leu-----IleAsnGlnAsnCysAspSerAlaArgThrSerAspAspArgLeu 112
|||||
Db 282 CTGAGAGTTCAGGAGTGTGATTCACAGTGCACATCTCCATCTCCACAGTAA----- 229
|||||

QY 113 CysArgSer 115
|||||

Db 228 TGTGAAGC 220

RESULT 15
AI081371/c 454 bp mRNA linear EST 13-AUG-1998
LOCUS AI081371.1
DEFINITION ox76910.x1 Soares_NHMPu.S1 Homo sapiens cDNA clone IMAGE:1662306

ACCESSION 3', mRNA sequence.
VERSION AI081371
KEYWORDS AI081371.1 GI:3418163
SOURCE EST.
ORGANISM human.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 454)
TITLE NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapb-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 355 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 392.
Location/Qualifiers
1. 454
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1662306"
/issue_type="Soares_NHMPu.S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: p773D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2Nbhm, pregnant uterus
Nbhpv, and fetal heart Nbhm19v) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479." 1 others

BASE COUNT 114 a 107 c 102 g 131 t

ORIGIN

Alignment Scores:
Pred. No.: 5.62e-08 Length: 454
Score: 170.50 Matches: 36
Percent Similarity: 57.69% Conservative: 9
Best Local Similarity: 46.15% Mismatches: 20
Query Match: 25.52% Indels: 13
DB: Gaps: 4

US-09-092-297-17 (1-117) x AI081371 (1-454)

QY 49 ArgCysGlyAspLysPheTyrAspProLeuGlnHisCysCysTyrAspAspAlaValVal 68
|||||
Db 451 AGGTGTGAGAGACAGACTACACCCCTGGAGCAGTGTTCATTAAGAGCCATGTG 392
|||||

QY 69 ProLeuAlaArgThrGlnThrCysGly---AsnCysThrPheArgValCysPheGln 87
:::|||||
Db 391 TCCCTGAGCGAGACCCGCCAATGTGTTGCCCTGCACCTCTGCTTGAAGCTC 332
|||||

QY 88 CysCys-----ProTyrPheMetValLysIleu----- 97
|||||

Db 331 TGTGTCTTGTATCTTCTTGGCTTCACAAAGATTTTGTGAAGCTGAAGTTCAGGCT 272
|||||

QY 98 IleAsnGlnAsnCysAspSerAlaArgThrSerAspAspArgLeuCysArgSer 115
:::|||||
Db 271 GTGAATTCAGTCCAGTCCATCTCCATCTCCAGTAA-----TGTGAAGC 224
|||||

Search completed: November 3, 2002, 05:01:19
Job time: 1693 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 3, 2002, 01:14:46 ; Search time 46 Seconds

(without alignments)
624.763 Million cell updates/sec

Title: US-09-092-297-17

Perfect score: 668
Sequence: 1 PQQPRAMPKRCIVAVFAI.....INONCDARTSDRLCRSVS 117

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Egapop 10.0 , Egapext 0.5	
Dgapop 6.0 , Dgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DEV=x1h
-O/cgn2_1/USPTO.SP001/US09092297/rnat.29102002.093920.21691/app.query.fasta_1.263
-DB=Issued Patents.NA -OFMT=fastp -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdt
-LIST=45 -DOCALLGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXPU=6 -ICPU=3
-USER=US09092297.ecgn.1.1.13.gnumat.29102002.093920.21691 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents.NA.*

1:	/cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2:	/cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3:	/cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4:	/cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5:	/cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
6:	/cgn2_6/ptodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78.5	11.8	2485	1	US-08-424-424B-1
2	78.5	11.8	2486	5	PCR-US94-05363A-1
3	75.5	11.3	4403765	4	US-09-103-840A-2
4	74	11.1	14042	4	US-08-652-877-85
5	74	11.1	14044	4	US-08-652-877-89
6	74	11.1	14080	4	US-08-652-877-89
7	74	11.1	14083	4	US-08-476-515A-83
8	74	11.1	14083	4	US-08-652-877-83
9	73	10.9	1194	1	US-08-447-702-4
10	73	10.9	1194	1	US-08-447-702-4
11	73	10.9	5312	4	US-08-485-355B-39
12	73	10.9	5312	4	US-08-485-355B-41

Result No.	Score	Query Match	Length	ID	Description
13	73	10.9	5312	4	US-08-485-355B-43
14	73	10.9	5368	4	US-08-485-355B-45
15	71	10.6	1254	1	US-08-313-553-1
16	71	10.6	1254	3	US-08-767-993-1
17	71	10.6	1265	4	US-09-020-956-173
18	71	10.6	1265	4	US-09-030-607-173
19	71	10.6	1265	4	US-09-439-313-173
20	71	10.6	1557	1	US-08-385-229-3
21	71	10.6	1641	1	US-08-385-229-1
22	71	10.6	1641	2	US-08-650-000-1
23	71	10.6	1641	6	5395760-1
24	71	10.6	2147	1	US-08-313-553-14
25	71	10.6	2147	3	US-08-767-993-14
26	71	10.6	3713	1	US-08-330-537-3
27	71	10.6	3713	1	US-08-330-537-4
28	71	10.6	4411529	4	US-09-103-840A-1
29	70.5	10.6	702	4	US-09-357-251-15
30	70.5	10.6	996	2	US-08-624-650-2
31	70.5	10.6	4722	4	US-08-979-608A-14
32	70	10.5	1560	3	US-08-629-643A-3
33	70	10.5	1560	4	US-09-155-884-3
34	70	10.5	6412	4	US-08-652-877-17
35	70	10.5	6412	4	US-08-476-515A-17
36	70	10.5	8982	3	US-08-976-255-5
37	69.5	10.4	430	4	US-08-905-223-231
38	69.5	10.4	1202	3	US-08-728-603-16
39	69.5	10.4	2934	3	US-09-149-934-2
40	69.5	10.4	32207	2	US-08-770-379-20
41	69.5	10.4	32207	4	US-08-757-669A-20
42	69.5	10.4	32207	4	US-09-230-371A-20
43	69	10.3	1545	4	US-08-900-117A-2
44	69	10.3	2224	4	US-08-477-347-2
45	69	10.3	2224	4	US-08-476-862-1

ALIGNMENTS

RESULT 1
US-08-424-424B-1
Sequence 1, Application US/08424424B
Patent No. 5758854
GENERAL INFORMATION:
APPLICANT: LI, ET AL.
TITLE OF INVENTION: Neurotransmitter Transporter
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESS: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,424B
FILING DATE: APRIL 21, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05363
FILING DATE: MAY 25, 1996
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-308
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 2485 BASE PAIRS
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
US-08-424-424B-1

Alignment Scores:
Pred. No.: 6.99 Length: 2485
Score: 78.50 Matches: 30
Percent Similarity: 35.58% Conservative: 7
Best Local Similarity: 28.85% Mismatches: 39
Query Match: 11.75% Indels: 28
DB: 1 Gaps: 6

US-09-092-297-17 (1-117) x US-08-424-424B-1 (1-2485)

QY 5 PROARGALAMETALAPROARGLYCYSLILEVALALVALPHEALALE-PHECYSILESE 24
DB 1739 CCCAAGAGATGTTACAGTGGCTGCTGTTCATTACATTCCTCGTGGAGACTGT----- 1793
QY 24 RARGLEUCYSSERHISGLYALAPROYALAPROMETHPROTYLLEUWELLEUCY 44
DB 1794 -----TGTTCTCCAGCGCTCCGGAACACTACTTGTGCACCATGTTCCATGACTAC 1843
QY 44 SGLNPROHLSLYSARGCYGLYASP----- 52
DB 1844 TCAGCA-----CGCTGCACACTCATCATGTCATCCTTGAGAACATCGCTGTGGCC 1897
QY 53 -LYSPHEIYRAPPROLEUGLNHISYS-----CYSTYRASPAPALAYA 67
DB 1898 TGGATTATGAGCCCAAGAGTTGATGACGAGCTGACGAGATGCTG-----GCTTC 1951
QY 67 IVALPROLEUALARGTGHTGTHCYSGLYASNCYSTHRRHEARGVALCYSPHEGLUCL 87
DB 1952 CGCCCTACCGCTTCTATTCTACATGTGGAAGTTGCTGTCCACTAGCATG-----CT 2008
QY 87 nCysCysPro 90
DB 2009 GTGCTCACCA 2018

RESULT 2
PCT-US94-05363A-1
Sequence 1, Application PC/TUS9405363A
GENERAL INFORMATION:
APPLICANT: LI, ET AL.
TITLE OF INVENTION: Neurotransmitter Transporter
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05363A
FILING DATE: SUBMITTED HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-118

TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2486 BASE PAIRS
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
PCR-US94-05363A-1

Alignment Scores:
Pred. No.: 6.99 Length: 2486
Score: 78.50 Matches: 30
Percent Similarity: 35.58% Conservative: 7
Best Local Similarity: 28.85% Mismatches: 39
Query Match: 11.75% Indels: 28
DB: 5 Gaps: 6

US-09-092-297-17 (1-117) x PCT-US94-05363A-1 (1-2486)

QY 5 PROARGALAMETALAPROARGLYCYSLILEVALALVALPHEALALE-PHECYSILESE 24
DB 1740 CCCAAGAGATGTTACAGTGGCTGCTGTTCATTACATTCCTCGTGGAGACTGT----- 1794
QY 24 RARGLEUCYSSERHISGLYALAPROYALAPROMETHPROTYLLEUWELLEUCY 44
DB 1795 -----TGTTCTCCAGCGCTCCGGAACACTACTTGTGCACCATGTTCCATGACTAC 1844
QY 44 SGLNPROHLSLYSARGCYGLYASP----- 52
DB 1845 TCAGCA-----CGCTGCACACTCATCATGTCATCCTTGAGAACATCGCTGTGGCC 1898
QY 53 -LYSPHEIYRAPPROLEUGLNHISYS-----CYSTYRASPAPALAYA 67
DB 1899 TGGATTATGAGCCCAAGAGTTGATGACGAGCTGACGAGATGCTG-----GCTTC 1952
QY 67 IVALPROLEUALARGTGHTGTHCYSGLYASNCYSTHRRHEARGVALCYSPHEGLUCL 87
DB 1953 CGCCCTACCGCTTCTATTCTACATGTGGAAGTTGCTGTCCACTAGCATG-----CT 2009
QY 87 nCysCysPro 90
DB 2010 GTGCTCACCA 2019

RESULT 3
US-09-103-840A-2
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103.840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

Alignment Scores:

Pred. No.: ..	3,77e+05		Length:	4403765
Score:	75.50		Matches:	23
Percent Similarity:	42.86%		Conservative:	10
Best Local Similarity:	29.87%		Mismatches:	25
Query Match:	11.30%		Indels:	19
DB:	4		Gaps:	5

US-09-092-297-17 (1-117) x US-09-103-840A-2 (1-4403765)

Oy 1 ProleugInProProArGAlaMetaIAProArG-----GLCySILLeValaIalValPhe 18
 |||::: :|||::| |
Db 4256986 CCgATTCAtGGCGGGGGATCCCTGTGCCTCCGGGTtTtGCCAGCCTCACACCCACACCcCGG 4257045
 ::||||| ||||| ||| |||:
Oy 19 AlaietheCySlISerArgLeuLeuCysSerHslsgYalaIProValaIAPrometthr 38
 ::||||| ||||| ||| |||:
Db 4257046 GTgtTGTTttGCCACACAGCACTGGCATGCACCCAGCGGTG----- 4257087
Oy 39 ProTYrLeuMeLeuCYGlNProHisLys---ArgCySGlyAspLyPhetYraSPPro 57
 ---||| ttcCACATTtccACcCCAGCcGtGTGTGTctccc---CACAttCCG 4257129
Db 4257088 -----

Oy 58 Leu-----GlnHlScycscytTyraSPasPaIalValaIProLeu 70
 ||| :|||::| |||::: ||||| |||
Db 4257130 TTggTfTCACACCGCCGAATTGCTGCATGCCGCGCCCGAGATCGCACCGTtg 4257180

RESULT 4
US-08-652-877-85
Sequence 85, Application US/08652877
Patent No. 6187548
GENERAL INFORMATION:
APPLICANT: Akerstrom, Goran
APPLICANT: Juhlin, Claes
APPLICANT: Rask, Lars
APPLICANT: Cronley, Gregg R.
APPLICANT: Morse, Clarence C.
APPLICANT: Murray, Edward M.
APPLICANT: Hjaln, Goran
TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
TITLE OF INVENTION: Thereof and DNA Encoding Same
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorel Inc.
STREET: 500 Arcola Rd., 3c43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426-0107
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.5.1
SOFTWARE: Word 6.0 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,877
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15203
FILING DATE: 22-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,836
FILING DATE: 23-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,314
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky, Martin
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: AI355E-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-454-3816
TELEFAX: 610-454-3808
INFORMATION FOR SEQ ID NO: 85:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 14042 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEITICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: Homo sapiens
ORGANISM: Homo sapiens
TISSUE TYPE: Placenta
FEATURE:
NAME/KEY: CDS
LOCATION: 68..14035
US-08-652-877-85

Alignment Scores:
Pred. No.: 270
Score: 74.00
Percent Similarity: 30.71%
Best Local Similarity: 25.20%
Query Match: 11.08%
DB: 4

US-09-092-297-17 (1-117) x US-08-652-877-85 (1-14042)
OY 28 CysSerHisGlyAlaProValAlaProMetThrProTyrLeuMetLeuCysGlnProHis 47
Db 9065 TGTGGTATGCGACGCGTGTATC-----CCAAAGATATTGACGGTGTACCGCGCAC 9112
OY 48 LysArgCysGlyAspLysPheTyrAspProLeuGlnHisCysGlyTyrAspAlaVal 67
Db 9113 AATGACTGTGGTGAC-----TATAGGACGAGAGGGGCGTGTATAC----- 9154
OY 68 ValProLeuAlaGlyThrGlnThrCys-----GlyAsn 78
Db 9155 -----CAGAGCTTCCCAACAGAAATCACTTTACCTGTACAGAACGGCGC 9196
OY 79 Cys-----ThrPheArgValCysPheGlnGln----- 87
Db 9197 TGCATTGTGTAACCTTC---GTCTGTGATGAGATATGACTGTGGAGACGGATCTGAT 9253
OY 88 -----CysCysProThrPhe----- 93
Db 9254 GAGCTATGCACTGTGACCAACCCAGAACCCAGCTGTCCACCTCAGAGATCTCAAGT 9313
OY 94 -----MetValLysLeuIleAsnGlnAsnCysAspSerAlaArg 106
Db 9314 GACAATGGCGCTGCATCGAGATGATGAAGAACTGCAACACCACTGATGACTGTTTGAC 9373
OY 107 ThrSerAspArgLeuGly 113
Db 9374 AACAGCGATGAGAAAGCTGT 9394

RESULT 5
US-08-652-877-89
Sequence 89, Application US/08652877
Patent No. 6187548
GENERAL INFORMATION:
APPLICANT: Akertstrom, Goran
APPLICANT: Juhlin, Claes
APPLICANT: Raak, Lars
APPLICANT: Crumley, Gregg R.
APPLICANT: Morse, Clarence C.
APPLICANT: Murray, Edward M.
APPLICANT: Hyalm, Goran
TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
TITLE OF INVENTION: Therapeutic and DNA Encoding Same
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Atcoia Rd., 3C43
CITY: Collegeville

```

```

STATE: PA
COUNTRY: USA
ZIP: 19426-0107
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.5.1
SOFTWARE: Word 6.0 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,877
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15203
FILING DATE: 22-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,836
FILING DATE: 23-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,314
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky, Martin
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: A1355E-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-454-3816
TELEFAX: 610-454-3808
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Parathyroid
FEATURE:
NAME/KEY: CDS
LOCATION: 65..14032
US-08-652-877-89
Alignment Scores:
Pred. No.: 270 Length: 14044
Score: 74.00 Matches: 32
Percent Similarity: 30.718 Conservative: 7
Best Local Similarity: 25.208 Mismatches: 30
Query Match: 11.088 Indels: 58
DB: 4 Gaps: 8
US-09-092-297-17 (1-117) x US-08-652-877-89 (1-14044)
QY 28 CysSerHisGlyAlaProValAlaProMetThrProTyrLeuMetLeuGlnProHis 47
Db 9062 TGTGGTACGACGTGTATC-----CCAAAGATATTACGGTGTGACCGGCAC 9109
QY 48 LysArgCysGlyAspLysPheTyrAspProLeuGlnHisCysCysTyrAspAspAlaVal 67
Db 9110 AATGACGTGTGTAC-----TATAGCAGACGAGGGGCTGTATAC----- 9151
QY 68 ValProLeuAlaArgThrGlnThrCys-----GlyAsn 78
Db 9152 -----CAGACTTGGCAACAGATCATGTTACTCTGTGAGAACGGCGC 9193
QY 79 Cys-----ThrPheArgValCysPheGln----- 87
Db 9194 TGCATTAGTAACCTTC--GTCCTGATGAGATATGACTGTGAGACGATCTGAT 9250
QY 88 -----CysCysProThrPhe----- 93

```

```

Db 9251 GAGCTGATGCACCTGTGCGCACCCCGAACCCAGAGTGTCCAGCTCAGAGTTCAAGTGT 9310
QY 94 -----MetValIysIleuIleasnGlnIasnCysAspSerAlaArg 106
Db 9311 GACATAGCGCCCTGCATGCATGATGATGAACCTGTGCAACCACTAGATGATGTTGGAC 9370
QY 107 ThrSerAspArgLeuCys 113
Db 9371 AACAGCATGAGAAAGCTGT 9391
RESULT 6
US-08-652-877-87
Sequence 87, Application US/08652877
Patent No. 6187548
GENERAL INFORMATION:
APPLICANT: Aketstrom, Goran
APPLICANT: Junlin, Claes
APPLICANT: Raak, Lars
APPLICANT: Crumley, Gregg R.
APPLICANT: Morse, Clarence C.
APPLICANT: Murray, Edward M.
APPLICANT: Hjaln, Goran
TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
TITLE OF INVENTION: Theeof and DNA Encoding Same
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Rd., 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426-0107
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.5.1
SOFTWARE: Word 6.0 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,877
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15203
FILING DATE: 22-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,836
FILING DATE: 23-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,314
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky, Martin
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: A1355E-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-454-3816
TELEFAX: 610-454-3808
INFORMATION FOR SEQ ID NO: 87:
SEQUENCE CHARACTERISTICS:
LENGTH: 14080 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Kidney
FEATURE:
NAME/KEY: CDS
LOCATION: 105..14072
US-08-652-877-87

```

Alignment Scores:

Pred. No.: 271 Length: 14080
 Score: 74.00 Matches: 32
 Percent Similarity: 30.71% Conservative: 7
 Best Local Similarity: 25.20% Mismatches: 30
 Query Match: 11.08% Indels: 58
 DB: 4 Gaps: 8

US-09-092-297-17 (1-117) x US-08-652-877-87 (1-14080)

QY 28 CysSerHisGlyAlaProValAlaPromethrProTyrLeuMetLeuGlyGlnProHis 47
 DB 9102 TGTGGTTACGACGAGTGTATC-----CCAAAGATATTGAGTGTGACCGGCAC 9149
 QY 48 LysArgGlyGlyAspLysPheTyrAspProLeuGlnHisCysCysTyrAspAlaVal 67
 DB 9150 AATGACTGTGTGAC-----TATAGCGACGAGAGGGGCTGCTTATAC----- 9191
 QY 68 ValProLeuAlaArgThrGlnThrCys-----GlyAsn 78
 DB 9192 -----CAGACTTGCCACACAGATCAGTTTACTGTCAGAACGGGCGC 9233
 QY 79 Cys-----ThrPheArgValCysPheGluGln----- 87
 DB 9234 TGCATTAGTAAACCTTC---GTCTGTGATGAGGATATGACTGTGAGACGATCTGAT 9290
 QY 88 -----CysCysProTyrThrPhe----- 93
 DB 9291 GAGCTGATGACACTGTGTCACACCCAGAACCCAGCGTGTCCACCTCAGATTCAGTGT 9350
 QY 94 -----MetValLysLeuIleAsnGlnAsnCysAspSerAlaArg 106
 DB 9351 GACAAATGGCGCTGCATCGAGATGATGTAACCTGTCAACACCACTAGACTGTTTGGAC 9410
 QY 107 ThrSerAspArgLeuGlyCys 113
 DB 9411 AACAGCGATGAGAAAGCTGT 9431

RESULT 7

US-08-476-515A-83
 ; Sequence 83, Application US/08476515A
 ; Patent No. 6239270
 ; GENERAL INFORMATION:
 ; APPLICANT: Aketstrom, Goran
 ; APPLICANT: Junlin, Claes
 ; APPLICANT: Raak, Lars
 ; APPLICANT: Crumley, Gregg R.
 ; APPLICANT: Morse, Clarence C.
 ; APPLICANT: Murray, Edward M.
 ; APPLICANT: Halm, Goran
 ; TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
 ; TITLE OF INVENTION: Thereof and DNA Encoding Same
 ; NUMBER OF SEQUENCES: 84
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Martin Savitzky
 ; STREET: Rhone-Poulenc Rorer Inc., 500 Arcoia Rd.,
 ; STREET: 3C43,
 ; CITY: Collegeville
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19426-0107
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Compaq PC
 ; OPERATING SYSTEM: Windows 95
 ; SOFTWARE: Word 7.0 (Patentlin)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/476,515A
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/344,836

FILING DATE: 23-NOV-1994

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/SE94/00483
 ; FILING DATE: 24-MAY-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: SE 9301764-8
 ; FILING DATE: 24-MAY-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Savitzky, Martin
 ; REGISTRATION NUMBER: 29,699
 ; REFERENCE/DOCKET NUMBER: A1355D
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 610-454-3816
 ; TELEFAX: 610-454-3808
 ; INFORMATION FOR SEQ ID NO: 83:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 14083 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 107..14071
 ; US-08-476-515A-83

Alignment Scores:

Pred. No.: 271 Length: 14083
 Score: 74.00 Matches: 32
 Percent Similarity: 30.71% Conservative: 7
 Best Local Similarity: 25.20% Mismatches: 30
 Query Match: 11.08% Indels: 58
 DB: 4 Gaps: 8

US-09-092-297-17 (1-117) x US-08-476-515A-83 (1-14083)

QY 28 CysSerHisGlyAlaProValAlaPromethrProTyrLeuMetLeuGlyGlnProHis 47
 DB 9104 TGTGGTTACGACGAGTGTATC-----CCAAAGATATTGAGTGTGACCGGCAC 9151
 QY 48 LysArgGlyGlyAspLysPheTyrAspProLeuGlnHisCysCysTyrAspAlaVal 67
 DB 9152 AATGACTGTGTGAC-----TATAGCGACGAGAGGGGCTGCTTATAC----- 9193
 QY 68 ValProLeuAlaArgThrGlnThrCys-----GlyAsn 78
 DB 9194 -----CAGACTTGCCACACAGATCAGTTTACTGTCAGAACGGGCGC 9235
 QY 79 Cys-----ThrPheArgValCysPheGluGln----- 87
 DB 9236 TGCATTAGTAAACCTTC---GTCTGTGATGAGGATATGACTGTGAGACGATCTGAT 9292
 QY 88 -----CysCysProTyrThrPhe----- 93
 DB 9293 GAGCTGATGACACTGTGTCACACCCAGAACCCAGCGTGTCCACCTCAGATTCAGTGT 9352
 QY 94 -----MetValLysLeuIleAsnGlnAsnCysAspSerAlaArg 106
 DB 9353 GACAAATGGCGCTGCATCGAGATGATGTAACCTGTCAACCACTAGACTGTTTGGAC 9412
 QY 107 ThrSerAspArgLeuGlyCys 113
 DB 9413 AACAGCGATGAGAAAGCTGT 9433

RESULT 8

US-08-652-877-83
 ; Sequence 83, Application US/08652877
 ; Patent No. 6187548
 ; GENERAL INFORMATION:
 ; APPLICANT: Aketstrom, Goran
 ; APPLICANT: Junlin, Claes

APPLICANT: Rask, Lars
APPLICANT: Crumley, Gregg R.
APPLICANT: Morse, Clarence C.
APPLICANT: Murray, Edward M.
APPLICANT: Hjaltn, Goran
TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
TITLE OF INVENTION: thereof and DNA Encoding Same
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Rd., 3043
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426-0107
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.5.1
SOFTWARE: Word 6.0 (Patentlin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,877
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15203
FILING DATE: 22-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,836
FILING DATE: 23-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,314
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky, Martin
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: A1355E-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-454-3816
TELEFAX: 610-454-3808
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 14086 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 107..14074
US-08-652-877-83
US-09-092-297-17 (1-117) x US-08-652-877-83 (1-14086)
Alignment Scores:
Pred. No.: 271 Length: 14086
Score: 74.00 Matches: 32
Percent Similarity: 30.71% Conservative: 7
Best Local Similarity: 25.20% Mismatches: 30
Query Match: 11.08% Indels: 58
Gaps: 8
DB: 4
QY 28 CysSerHisGlyAlaProValAlaProMetThrProTyrLeuMetLeuCysGlnProHis 47
DB 9104 TGTGTTACGAGCTGTATC-----CCAAAGATATTACAGTGTGACCGGCAC 9151
QY 48 LysArgCysGlyAspLysPheTyrAspProLeuGlnHisCysCysTyrAspAlaVal 67
DB 9152 AATACAGTGTGTATC-----TATAGGACGAGAGGGGCTGCTTATAC----- 9193

QY 68 ValProLeuAlaArgThrGlnThrCys-----GlyAsn 78
DB 9194 -----CAGACTTGCCACAGATCAGTTTACTCTGCACAGCGGCGC 9235
QY 79 Cys-----ThrPheArgValCysPheGluIn----- 87
DB 9236 TGCATTAGTAACCTTC---GTCTGTATGAGATATGACTGTGAGACGATCTGAT 9292
QY 88 -----CysCysProTyrPhe----- 93
DB 9293 GAGCTGATGCACCTGTGCGCACACCCAGAACCCAGTGTCCACCTCAGAGTTCAAGTGT 9332
QY 94 -----MetValLysLeuIleAsnGlnAsnCysAspSerAlaArg 106
DB 9353 GACAATGGCGGCTCATGATGATGATGAATCTGTCAACCCATGATGACTGTTTGAC 9412
QY 107 ThrSerAspArgLeuCys 113
DB 9413 AACAGCATGAGAAAGCTGT 9433
RESULT 9
US-08-447-702-4
Sequence 4, Application US/08447702
Patent No. 5629190
GENERAL INFORMATION:
APPLICANT: Petre, Dominique
APPLICANT: Cerbeaud, Edith
APPLICANT: Levy-Schill, Sophie
APPLICANT: Crouzel, Joel
TITLE OF INVENTION: Polypeptides Possessing A Nitrlase
TITLE OF INVENTION: Activity, DNA Sequence Coding for Said Polypeptides,
Patent No. 5629190
TITLE OF INVENTION: Expression Cassettes and Host Microorganisms Enabling The
TITLE OF INVENTION: to be Obtained, and Method of Converting Nitrlases to Carb
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: P.O. Box 1404
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentlin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,702
FILING DATE: 23-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/194,588
FILING DATE: 10-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92-09-882
FILING DATE: 10-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 003025-019
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1194 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)

FILED DATE: 08-JUL-1993
APPLICATION NUMBER: AU PL4081/92
FILING DATE: 14-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Treacartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-58631-2/RFT/DSS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 5312 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 37..5148
SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-08-485-355B-39
Alignment Scores:
Pred. No.: 92.2 Length: 5312
Score: 73.00 Matches: 32
Percent Similarity: 37.40% Conservative: 14
Best Local Similarity: 26.02% Mismatches: 41
Query Match: 10.93% Indels: 36
Gaps: 7
US-09-092-297-17 (1-117) x US-08-485-355B-39 (1-5312)
QY 5 ProaTgAlaMetAlaProaTgAlaCysTleValAlaValAlaPheAlaIlePheCysTleSer 24
DB 2744 CCGTGGCGGCGTGGACACTGCGGGGGTTCGTCGACCTGCGGCGCTGCGCTCT 2685
QY 25 ArgLeu-----LeucYserHisGlyAlaProValAlaPheMetThrProTyr 40
DB 2684 TCACCTCGGTTGACAGCTGATCTCCAGGGGTCGAGGATGCTGAACACCTCGGCT 2625
QY 41 LeuMetLeu-----CysGlnProHisLysArgCysGlyAspLysPheTyrAsp--- 56
DB 2624 TTCGGCGTATGTTGAGTGTCTCTCAATGTCACCTGTCGGGTCGGGATGACAGGTGG 2565
QY 57 -----ProLeuGlnHisCysCysTyrAsp-----Asp 65
DB 2564 TTGGTGTGGCGCGGTATGCCGACTGAGAGGTGCGACTCTCTCAGCGAGAGCTTCTGCTCT 2505
QY 66 AlaValValProLeu-----AlaArgThrGlnThrCysGlyAsnCys 79
DB 2504 GCTGTGGAGCGGTGTATGATGACAGATGACAGACGAAAGGCGCTCGCTGCGCTGTC 2445
QY 80 Thr-----PheArgValCysPheGlnGln---Cys 88
DB 2444 ACAGTACGCGCGCCCTCAGCCCGCTGGCGGACGACTTCTCTCTCGGTGAGAGCAGCGCT 2385
QY 89 CysProTyr 91
DB 2384 TGGCGCTGG 2376
RESULT 12
US-08-485-355B-41/c
Sequence 41, Application US/08485355B
Patent No. 6177075
GENERAL INFORMATION:
APPLICANT: Christian, P. D., Gordon, K. H.J., Hanzlik, T. N.
TITLE OF INVENTION: Insect Viruses and Their Uses in
Protecting Plants
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP

STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,355B
FILING DATE: 07-Jun-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/440,522
FILING DATE: 12-MAY-1995
APPLICATION NUMBER: US/08/089,372
FILING DATE: 08-JUL-1993
APPLICATION NUMBER: AU PL4081/92
FILING DATE: 14-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Treacartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-58631-2/RFT/DSS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 5312 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 4218..4514
SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-08-485-355B-41
Alignment Scores:
Pred. No.: 92.2 Length: 5312
Score: 73.00 Matches: 32
Percent Similarity: 37.40% Conservative: 14
Best Local Similarity: 26.02% Mismatches: 41
Query Match: 10.93% Indels: 36
Gaps: 7
US-09-092-297-17 (1-117) x US-08-485-355B-41 (1-5312)
QY 5 ProaTgAlaMetAlaProaTgAlaCysTleValAlaValAlaPheAlaIlePheCysTleSer 24
DB 2744 CCGTGGCGGCGTGGACACTGCGGGGGTTCGTCGACCTGCGGCGCTGCGCTCT 2685
QY 25 ArgLeu-----LeucYserHisGlyAlaProValAlaPheMetThrProTyr 40
DB 2684 TCACCTCGGTTGACAGCTGATCTCCAGGGGTCGAGGATGCTGAACACCTCGGCT 2625
QY 41 LeuMetLeu-----CysGlnProHisLysArgCysGlyAspLysPheTyrAsp--- 56
DB 2624 TTCGGCGTATGTTGAGTGTCTCTCAATGTCACCTGTCGGGTCGGGATGACAGGTGG 2565
QY 57 -----ProLeuGlnHisCysCysTyrAsp-----Asp 65
DB 2564 TTGGTGTGGCGCGGTATGCCGACTGAGAGGTGCGACTCTCTCAGCGAGAGCTTCTGCTCT 2505
QY 66 AlaValValProLeu-----AlaArgThrGlnThrCysGlyAsnCys 79
DB 2504 GCTGTGGAGCGGTGTATGATGACAGATGACAGACGAAAGGCGCTCGCTGCGCTGTC 2445
QY 80 Thr-----PheArgValCysPheGlnGln---Cys 88

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 3, 2002, 03:14:04 ; Search time 210 Seconds

(without alignments)
956.567 Million cell updates/sec

Title: US-09-092-297-17

Perfect score: 117
Sequence: 1 PLDPPRAMPKRCIVAVFAL.....INONCSARISDRLCRSVS 117

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1736436 segs, 858457221 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3471288

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame-p2n.model -DEV=xlh
-Q/cgn2.1/USPTO.spool/US09092297/rnatc.29102002.094016.22405/app.query.fasta.1.263
-DB=N.Geneseq.032802 -OFMT=fastap -SUFFIX=olip2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09092297-@CEN.1.1.0-@rnatc.29102002.094016.22405 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-NARN_TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FAPOP=6 -FAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N.Geneseq.032802.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	117	100.0	762	AAZ24391	Human bladder tumor
2	117	100.0	763	AAZ24391	Urinary tract tiss
3	117	100.0	763	AAZ24391	Urinary tract tiss
4	117	100.0	771	AAA77668	Human PRO1274 CDNA
5	117	100.0	771	AAA77668	Human PRO1274 (UNC)
6	117	100.0	771	AAA77668	Human cDNA sequenc
7	117	100.0	771	AAA77668	Native sequence of
8	117	100.0	771	AAA77668	DNA encoding prote
9	117	100.0	771	AAA77668	Human bladder spec
10	116	99.1	748	AAZ98158	Human signal pepti
11	116	99.1	748	AAZ98158	Urinary tract tiss
12	116	99.1	748	AAZ98158	Urinary tract tiss
13	64	54.7	236	AAZ98158	DNA encoding novel
14	64	54.7	236	AAZ98158	Human breast cell
15	64	54.7	236	AAZ98158	Human foetal liver
16	64	54.7	236	AAZ98158	Probe #12918 for g
17	64	54.7	236	AAZ98158	Human bone marrow
18	64	54.7	236	AAZ98158	Probe #12228 for g
19	64	54.7	236	AAZ98158	Probe #16267 used
20	64	54.7	236	AAZ98158	Human breast cell
21	64	54.7	236	AAZ98158	Human foetal liver
22	64	54.7	236	AAZ98158	Probe #2990 for ge
23	64	54.7	236	AAZ98158	Human bone marrow
24	64	54.7	236	AAZ98158	Probe #2991 for ge
25	64	54.7	236	AAZ98158	Probe #1091 used t
26	64	54.7	236	AAZ98158	Arabidopsis thalia
27	64	54.7	236	AAZ98158	Arabidopsis thalia
28	64	54.7	236	AAZ98158	Arabidopsis thalia
29	64	54.7	236	AAZ98158	Arabidopsis thalia
30	64	54.7	236	AAZ98158	Arabidopsis thalia
31	64	54.7	236	AAZ98158	Drosophila melanog
32	64	54.7	236	AAZ98158	Drosophila melanog
33	64	54.7	236	AAZ98158	Drosophila melanog
34	64	54.7	236	AAZ98158	Drosophila melanog
35	64	54.7	236	AAZ98158	N. meningitidis pa
36	64	54.7	236	AAZ98158	N. meningitidis pa
37	64	54.7	236	AAZ98158	N. meningitidis pa
38	64	54.7	236	AAZ98158	N. meningitidis pa
39	64	54.7	236	AAZ98158	N. meningitidis pa
40	64	54.7	236	AAZ98158	N. meningitidis pa
41	64	54.7	236	AAZ98158	N. meningitidis pa
42	64	54.7	236	AAZ98158	N. meningitidis pa
43	64	54.7	236	AAZ98158	N. meningitidis pa
44	64	54.7	236	AAZ98158	N. meningitidis pa
45	64	54.7	236	AAZ98158	N. meningitidis pa

ALIGNMENTS

RESULT 1
ID AAZ24391 standard; cDNA; 762 BP.
AC AAZ24391;
XX 14-FEB-2000 (first entry)
DE Human bladder tumour CDNA library derived EST 3.
XX
XX
XX Expressed sequence tag; human; bladder; tumour; cancer; cytostatic;
KW treatment; gene therapy; EST; sa.
OS Homo sapiens.
XX
XX
XX DE19818619-A1.
XX
XX PD 28-OCT-1999.

	Pf	XX	21-APR-1998; 98DE-1018619.
	PR	XX	21-APR-1998; 98DE-1018619.
	PA	XX	(META-) METAGEN GES GENOMFORSCHUNG MBH.
	PI	XX	Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
	PT	DR	MPL 1999-612028/53.
	PP	XX	New nucleic acid sequences expressed in bladder tumor tissue, and derived polypeptides, for treatment of bladder tumor and identification of therapeutic agents
	PS	XX	Claim 3; Page 62; 132pp; German.
	CC	XX	This invention describes novel polypeptide fragments (I) and the polynucleotides (II) that encode them that are highly expressed in a human bladder tumour and which have cytostatic activity. (II) are used for recombinant expression of (I) and to isolate complete genes. (I) are used to identify agents suitable for treatment of bladder cancer, to directly treat this form of cancer (including expression from gene therapy vectors) or are used in a preparation for cancer treatment. (I) is also used for the generation of specific antibodies. (II) are identified by assembling ESTs (expressed sequence tags) from a particular tissue type before comparison of expression patterns. This allows a significantly longer fragment of the gene to be revealed, and therefore reduces the number of failures associated with the fact that ESTs from different libraries may represent different parts of the same unknown gene, distorting the estimated frequency of occurrence in a particular tissue. AA44360-743109 represent expressed sequence tag (EST) fragments isolated from a human bladder tumour cDNA library which encode the proteins represented in AA66143-Y66198.
	CC	XX	Sequence 762 BP; 162 A; 232 C; 194 G; 174 T; 0 other;
	SQ		
			Alignment Scores:
			Pred. No.: 5,62e-111 Length: 762
			Score: 117.00 Matches: 117
			Percent Similarity: 100.00% Conservative: 0
			Best Local Similarity: 100.00% Mismatches: 0
			Query Match: 100.00% Indels: 0
			DB: 20 Gaps: 0
			US-09-092-297-17 (1-117) x AA24391 (1-762)
OY	1	ProLeuGlnPrOPrOArGaAlMeLaLPArOGlYCylleValAlAvalPheNaIle 20	
Dd	3	CcATGCAGACCAACCACCGCATTGCCCTGGGTGATGTACTGTCTTGGCACATT 62	
OY	21	PheCySileSerArgLeuLeuCysSerHISglYAlAProVAlAlAPromethPrOTyr 40	
Dd	63	TTCGCAATCCACGGAGCTCCTCTCCACAAGGAGCCCCCATGACTACTCTTAC 122	
OY	41	LennetLeuGslnPrOHIsLysARgCYsgLYAsPLYSPHeTYrASPPROleuGlNHis 60	
Dd	123	CTGTAGTGtGtGCCAGCACCAAGAATGTGGGCAAACTTTACAGACCCCCTGCAGAAC 182	
OY	61	CyScSTYrASPaspAlaValAlVALProLeuAlAARGThGIInTHCySGLYASncSYthr 80	
Dd	183	TGTGTATATGATGATGGCGGCGTGGCCCTTGGCCAGAGACCAGACGTGTGGAAAAC TGACAC 242	
OY	81	PheArgValALcySpheGLnuGlcNYCSYPrtPrTHRPhemeTVallYSLeuIlEasngIN 100	
Dd	243	TTTCGAATCTGCTTTTAGAGCAGTGCTCCCTCGAGCTTCATGTGAACCTGATTAACAG 302	
OY	101	AsnCysApSeRIalargThrIserASPaspAYgleUCysAagseValser 117	
Dd	303	AACIGGCACATCAGCCCGGACCTCGGATGACAGGCTTTGTTCGAGTGTGACG 353	
		RESULT_2	
		AXX01686	

```

ID      AA01686 standard; DNA; 763 BP.
XX
AC      AA01686;
XX
DT      21-APR-1999 (first entry)
XX
DE      Urinary tract tissue library BL172 gene full length clone 15548381H.
XX
XX      BL172; urinary tract; marker; cancer; recombinant; human; ss.
XX
OS      Homo sapiens.
XX
PN      M09855656-A1.
XX
PD      10-DEC-1998.
XX
PF      05-JUN-1998; 98WO-US11693.
XX
PR      05-JUN-1997; 97US-0869579.
XX
PA      (ABBO ) ABBOTT LAB.
XX
PI      Billing-medel PA, Cohen M, Colpitts TL, Friedman PN;
PI      Gordon J, Granadosen, Hodges SC, Klass MR, Kratochvil JD;
PI      Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;
XX
DR      WPI: 1999-045802/04.
XX
PT      New purified polynucleotide BL172 derivatives and encoded
PT      polynucleotides - useful in the identification of markers, which are
PT      indicative of urinary tract diseases or conditions
XX
XX      Claim 1; Page 89-90; 114pp: English.
XX
XX      This represents a full-length sequence of a clone specific for urinary
XX      tract tissue library BL172. The BL172 polynucleotides (AA01683-87) are
XX      used to detect target BL172 polynucleotides and BL172 mRNA in a test
XX      polypeptides are used to detect BL172 antigens/antibodies in a test
XX      sample. The identification of certain markers in these methods are
XX      indicative of the presence of urinary tract disease, especially cancer.
XX      Additionally, the polypeptides are used to detect antibodies that bind
XX      specifically to a BL172 epitope derived from a BL172 polypeptide. Host
XX      cells containing a recombinant expression vector comprising the BL172
XX      polynucleotide sequences are used to produce the polypeptides containing
XX      BL172 epitopes. The methods aid in diagnosis, detection, staging,
XX      monitoring, prognosis, in vivo imaging, prevention and treatment
XX      of diseases or conditions associated with BL172, especially urinary
XX      tract cancer.
XX
XX      Sequence 763 BP; 162 A; 231 C; 195 G; 175 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.: 5 63e-111 Length: 763
XX      Score: 117.00 Matches: 117
XX      Percent Similarity: 100.00% Conservative: 0
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 100.00% Indels: 0
XX      DB: 20 Gaps: 0
XX
US-09-092-297-17 (1-117) x AA01686 (1-763)
QY      1 ProLeuGlnProProArgAlaMetAlaProArgGlyCysIleValAlaAlaAlaAla 20
Db      3 CCACTGCACACACCCAGAGCCATGCGTCCCGAGGCGTCACTGCTTGGCATT 62
QY      21 PheGlyIleSerArgIleuLeuGlySerHisGlyAlaIleProValAlaProMetThrProTy 40
Db      63 TTCGCAATGCTCCAGGCTCTCTGCTCAACAGGAGGCCCATGTGCCCATGACTCTTAC 122
QY      41 LeuMetLeuGlnProHisIleValSerGlyValSerPheThrValAspProLeuGlnHis 60
Db      123 CTGATGCTGTGCCAGCCACACAGAGATGTGTGGGACAAAGTTCTTACAGACCCCTGCGACAC 182

```

Qy	61	CysCysTyrAspAspAlaValValProleuAlaArgThrInGlnThyCysGlyAsnGlyThr	80
Db	183		
Qy	81	PheArgValCysPheGluGlnCysCysProTrpThrPheMetValLeuIleAsnGln	100
Db	243		
Qy	101	AsnGlyAspSerAlaArgThrSerAspAspArgLeuGlyArgSerValSer	117
Db	303		
		AACTCGACGACTACGCCGACCTGGATGACAGCGCTTGTCCAGAGTTCAGC	353
RESULT 3			
ID	AAx01687	standard; DNA; 763 BP.	
AC	AAx01687;		
XX			
DT	21-APR-1999	(first entry)	
XX			
DE	Urinary tract tissue library BL172 gene consensus sequence.		
XX			
KM	BL172; urinary tract; marker; cancer; recombinant; human; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	3..356	
FT		/*tag= a	
FT		/product= "BL172 polypeptide"	
FT		/note= "the start codon is not indicated"	
XX			
PN	W09855656-A1.		
XX			
PD	10-DEC-1998.		
XX			
PF	05-JUN-1998;	98W0-US11693.	
XX			
PR	05-JUN-1997;	97US-0869579.	
XX			
PA	(ABBO) ABBOTT LAB.		
XX			
PI	Billig-medel PA, Cohen M, Colpits TL, Friedman PN;		
XX	Gordon J, Grandosien, Hodges SC, Klass MR, Kratochvil JD;		
PI	Roberts-Rapp L, Russell JC, Strome SD, Yu H;		
XX			
DR	WPI: 1999-045802/04.		
XX			
DR	P-FSDB: AAW92252.		
XX			
PT	New purified polynucleotide BL172 derivatives and encoded		
XX	polypeptides - useful in the identification of markers, which are		
PT	indicative of urinary tract diseases or conditions		
XX			
PS	Claim 1; Page 90; 114pp; English.		
XX			
CC	This represents a consensus sequence of a clone specific for urinary		
CC	tract tissue library BL172. The BL172 polynucleotides (AAx01683-87) are		
CC	used to detect target BL172 polynucleotides and BL172 mRNA and the BL172		
CC	polypeptides are used to detect BL172 antigens/antibodies in a test		
CC	sample. The identification of certain markers in these methods are		
CC	indicative of the presence of urinary tract disease, especially cancer.		
CC	Additionally, the polypeptides are used to detect antibodies that bind		
CC	specifically to a BL172 epitope derived from a BL172 polypeptide. Host		
CC	cells containing a recombinant expression vector comprising the BL172		
CC	polynucleotide sequences are used to produce the polypeptides containing		
CC	BL172 epitopes. The methods aid in diagnosis, detection, staging,		
CC	monitoring, prognostication, in vivo imaging, prevention and treatment		
CC	of diseases or conditions associated with BL172, especially urinary		
XX	tract cancer.		
XX			
SQ	Sequence 763 BP; 162 A; 231 C; 194 G; 175 T; 1 other;		
Alignment Scores:			

Pred. No.:	5,63e-111	Length:	763
Score:	117.00	Matches:	117
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	20	Gaps:	0
US-09-092-297-17 (1-117) x AAX01687 (1-763)			
Qy 1	ProLeuGlnGInProPArgAlaMetAlaProArgGlyCysIleValAlaValAlaPheAlaIle 20		
Db 3	CCACTGCACACACCCAGAGCCATGGCTCCCCGAGGCTGATCTAGCTGCTTGGCATT 62		
Qy 21	PheCysIleSerAlaGleuLeuCysSerHisGlyAlaProValAlaPromethrProtyr 40		
Db 63	TTTCGACATCTCCAGAGCTCTCTGTCTACAGAGAGCCCAATGGCCCCCATGACTCTTAC 122		
Qy 41	LeuMetLeuCysGlnProHisLysAlaGlyAspIlyAspIlyAspPheTyraSpProLeuGlnHis 60		
Db 123	CTGATGCTGCTGCCAGCCACACAGAGATGTGGGGAGCAAGTTCTACAGACCCCTGGCAGCAC 182		
Qy 61	CysCysTyraSpAspAlaValAlaValProLeuAlaArgThrGlnhrCysGlyAsnCysThr 80		
Db 183	TGTGCTCATGTGATGATGCCGTGCTGCCCTTGGCCAGAGCCAGACGCTGTGAAACTGCACAC 242		
Qy 81	PheArgValCysPheGluGlnCysCysProTyrThrPheMetValLysLeuIleAsnGln 100		
Db 243	TTTCAGAGCTCTCTTGTGAGCAGCTGCTGCCCTTGACCTTCAATGTGTAAGCTGATTAACCAAG 302		
Qy 101	AsnCysAspSerAlaArgThrSerAspAspArgLeuCysArgSerValSer 117		
Db 303	AACGCACTCAGCCCGGACCTCGGATGACAGAGCTTTGTCGAGTGTACAGC 353		
RESULT 4			
AAA77668			
ID AAA77668 standard; cDNA; 771 BP.			
XX			
AC AAA77668;			
XX			
DE 07-NOV-2000 (first entry)			
XX			
DE Human PRO1274 cDNA sequence SEQ ID NO:196.			
XX			
DE Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;			
KW diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;			
KW angiogenic; proliferative; cardiac; cardiovascular; antiatherosclerotic;			
KW cytosolic; gene therapy; vaccine; ss.			
XX			
OS Homo sapiens.			
XX			
FN MO200032221-A2.			
XX			
XX 08-JUN-2000.			
XX			
PF 30-NOV-1999; 99WO-US28313.			
XX			
PR 01-DEC-1998; 98WO-US25108.			
PR 16-DEC-1998; 98US-0112850.			
PR 12-JAN-1999; 99US-0115554.			
PR 08-MAR-1999; 99WO-US05028.			
PR 12-MAR-1999; 99US-0123957.			
PR 28-APR-1999; 99US-0131445.			
PR 14-MAY-1999; 99US-0134287.			
PR 02-JUN-1999; 99WO-US12252.			
PR 23-JUN-1999; 99US-0141037.			
PR 20-JUL-1999; 99US-0144758.			
PR 26-JUL-1999; 99US-0145698.			
PR 01-SEP-1999; 99WO-US20111.			
PR 08-SEP-1999; 99WO-US20594.			
PR 13-SEP-1999; 99WO-US20944.			
PR 15-SEP-1999; 99WO-US21090.			
PR 15-SEP-1999; 99WO-US21547.			
PR 05-OCT-1999; 99WO-US23089.			

PR 29-OCT-1999; 99US-0162506.
XX
XX (GETH) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ, Goddard A;
PI Godowski PJ, Gurney AL, Klein RD, Kuo SS, Paoni NF, Smith V;
PI Matanabe CK, Williams PM, Wood WT.
XX
XX WPI; 2000-412154/35.
DR P-PSDB; AAB24425.
XX
XX Nucleic acids encoding PRO polypeptides useful for preventing
PT diagnosing and treating disorders a cardiovascular, endothelial or
PT angiogenic disorders in mammals -
XX
XX Claim 61; Fig 75; 315pp; English.
XX
XX The present invention describes nucleic acids encoding PRO polypeptides
CC useful for preventing, diagnosing and treating disorders a
CC cardiovascular, endothelial or angiogenic disorder in mammals by
CC modulating cell proliferation, angiogenesis and cardiovascularisation,
CC and for identifying agonists and antagonists of these processes. The
CC nucleic acids and the proteins they encode may be used in the
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate PRO expression such as cardiovascular, endothelial or
CC angiogenic disorders in mammals (e.g. atherosclerosis, cancers and
CC cardiac hypertrophy). For example, the nucleic acids (NCS) and vectors
CC containing them and the PRO polypeptide may be used to treat disorders
CC associated with decreased PRO expression. AAA77510 to AAA77721 and
CC AAB24388 to AAB24435 represent nucleotide and protein sequences used in
CC the exemplification of the present invention.
XX
SQ Sequence 771 BP; 169 A; 231 C; 195 G; 176 T; 0 other;

Alignment Scores:
Pred. No.: 5,68e-111 Length: 771
Score: 117.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-09-092-297-17. (1-117) x AAA77668 (1-771)
QY 1 PROLeuGlnProProAlaGlnAlaMetAlaProArgGlyCysIleValAlaValAlaIle 20
DB 3 CCAGTGCACACACAGAGCATGCTCCCGAGGCTGATGCTGATGCTGCTTGCATT 62
QY 21 PhcCysIleSerArgLeuLeuGlySerHisGlyAlaProValAlaProMetThrProTyr 40
DB 63 TTCTGATCTCCAGGCTCTCTGCTCACAAGAGCCCAAGTGGCCCACTGATCTCTTAC 122
QY 41 LeuMetLeuGlnProHisLysArgGlyAspLysPheThrAspProLeuGlnHis 60
DB 123 CTGATCTGTGCGACGACACAGAGATGTGGGACAAAGTTCTACACACCCCTGCAGCAC 182
QY 61 CysCysTyrAspAspAlaValAlaProLeuAlaArgThrGlnThrcysGlyAsnCysThr 80
DB 183 TGTTCCTATGATGATGATCCCTGCGCCCTGGCCAGAGACCAAGCGTGTGAAGTCAAC 242
QY 81 PheArgValCysPheGlnLincysCysProTrrPthPheMetValLysLeuIleAsnGln 100
DB 243 TTCAGAGTGTGCTTGAGAGAGTGTGCTCCCTGACCTTATGATGTAAGCTGATAACACAG 302
QY 101 AsnCysAspSerAlaArgThrSerAspAspArgLeuGlyAsrSerValSer 117
DB 303 AACTGGATCTACGCCGAGACCTCGATGACAGGCTTTGTCGAGTGTACG 353
RESULT 5
AAA37062
ID AAA37062 standard; cDNA; 771 BP.
XX
AC AAA37062;

XX
DT 08-AUG-2000 (first entry)
XX
XX Human PRO1274 (UN0644) cDNA sequence SFO ID NO:137.
DE
XX
XX Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
KW Transmembrane; secretion; immunoadhesion; pharmaceutical; screening;
KW ss.
XX
XX Homo sapiens.
XX
PN WO200012708-A2.
PD
XX
XX 09-MAR-2000.
XX
XX 01-SEP-1999; 99MO-US20111.
XX
XX 01-SEP-1998; 98US-0098716.
XX 01-SEP-1998; 98US-0098749.
XX 01-SEP-1998; 98US-0098750.
XX 02-SEP-1998; 98US-0098803.
XX 02-SEP-1998; 98US-0098821.
XX 02-SEP-1998; 98US-0098843.
XX 09-SEP-1998; 98US-0099536.
XX 09-SEP-1998; 98US-0099596.
XX 09-SEP-1998; 98US-0099598.
XX 09-SEP-1998; 98US-0099602.
XX 09-SEP-1998; 98US-0099642.
XX 10-SEP-1998; 98US-0099741.
XX 10-SEP-1998; 98US-0099754.
XX 10-SEP-1998; 98US-0099763.
XX 10-SEP-1998; 98US-0099792.
XX 10-SEP-1998; 98US-0099808.
XX 10-SEP-1998; 98US-0099812.
XX 10-SEP-1998; 98US-0099815.
XX 10-SEP-1998; 98US-0099816.
XX 15-SEP-1998; 98US-0100385.
XX 15-SEP-1998; 98US-0100388.
XX 15-SEP-1998; 98US-0100390.
XX 16-SEP-1998; 98US-0100584.
XX 16-SEP-1998; 98US-0100627.
XX 16-SEP-1998; 98US-0100661.
XX 16-SEP-1998; 98US-0100662.
XX 16-SEP-1998; 98US-0100664.
XX 17-SEP-1998; 98US-0100683.
XX 17-SEP-1998; 98US-0100684.
XX 17-SEP-1998; 98US-0100710.
XX 17-SEP-1998; 98US-0100711.
XX 17-SEP-1998; 98US-0100919.
XX 17-SEP-1998; 98US-0100930.
XX 18-SEP-1998; 98US-0100848.
XX 18-SEP-1998; 98US-0100849.
XX 18-SEP-1998; 98US-0101014.
XX 18-SEP-1998; 98US-0101068.
XX 18-SEP-1998; 98US-0101071.
XX 22-SEP-1998; 98US-0101279.
XX 23-SEP-1998; 98US-0101471.
XX 23-SEP-1998; 98US-0101472.
XX 23-SEP-1998; 98US-0101474.
XX 23-SEP-1998; 98US-0101475.
XX 23-SEP-1998; 98US-0101476.
XX 23-SEP-1998; 98US-0101477.
XX 23-SEP-1998; 98US-0101479.
XX 24-SEP-1998; 98US-0101738.
XX 24-SEP-1998; 98US-0101741.
XX 24-SEP-1998; 98US-0101743.
XX 24-SEP-1998; 98US-0101915.
XX 24-SEP-1998; 98US-0101916.
XX 29-SEP-1998; 98US-0102207.
XX 29-SEP-1998; 98US-0102240.
XX 29-SEP-1998; 98US-0102307.
XX 29-SEP-1998; 98US-0102330.
XX 29-SEP-1998; 98US-0102331.

PR	30-SEP-1998;	98US-0102484.
PR	30-SEP-1998;	98US-0102487.
PR	30-SEP-1998;	98US-0102570.
PR	30-SEP-1998;	98US-0102571.
PR	01-OCT-1998;	98US-0102684.
PR	01-OCT-1998;	98US-0102687.
PR	02-OCT-1998;	98US-0102965.
PR	06-OCT-1998;	98US-0103258.
PR	06-OCT-1998;	98US-0103449.
PR	07-OCT-1998;	98US-0103314.
PR	07-OCT-1998;	98US-0103315.
PR	07-OCT-1998;	98US-0103328.
PR	07-OCT-1998;	98US-0103395.
PR	07-OCT-1998;	98US-0103396.
PR	07-OCT-1998;	98US-0103401.
PR	08-OCT-1998;	98US-0103633.
PR	08-OCT-1998;	98US-0103678.
PR	08-OCT-1998;	98US-0103679.
PR	08-OCT-1998;	98US-0103711.
PR	14-OCT-1998;	98US-0104257.
PR	20-OCT-1998;	98US-0104987.
PR	20-OCT-1998;	98US-0105000.
PR	20-OCT-1998;	98US-0105002.
PR	21-OCT-1998;	98US-0105104.
PR	22-OCT-1998;	98US-0105169.
PR	22-OCT-1998;	98US-0105266.
PR	26-OCT-1998;	98US-0105693.
PR	26-OCT-1998;	98US-0105694.
PR	27-OCT-1998;	98US-0105807.
PR	27-OCT-1998;	98US-0105881.
PR	27-OCT-1998;	98US-0105882.
PR	27-OCT-1998;	98US-0106062.
PR	28-OCT-1998;	98US-0106023.
PR	28-OCT-1998;	98US-0106029.
PR	28-OCT-1998;	98US-0106030.
PR	28-OCT-1998;	98US-0106032.
PR	28-OCT-1998;	98US-0106033.
PR	28-OCT-1998;	98US-0106178.
PR	29-OCT-1998;	98US-0106248.
PR	29-OCT-1998;	98US-0106384.
PR	30-OCT-1998;	98US-0106500.
PR	30-OCT-1998;	98US-0106464.
PR	03-NOV-1998;	98US-0106856.
PR	03-NOV-1998;	98US-0106902.
PR	03-NOV-1998;	98US-0106905.
PR	03-NOV-1998;	98US-0106919.
PR	03-NOV-1998;	98US-0106932.
PR	10-NOV-1998;	98US-0106934.
PR	17-NOV-1998;	98US-0107783.
PR	17-NOV-1998;	98US-0108775.
PR	17-NOV-1998;	98US-0108779.
PR	17-NOV-1998;	98US-0108787.
PR	17-NOV-1998;	98US-0108788.
PR	17-NOV-1998;	98US-0108801.
PR	17-NOV-1998;	98US-0108802.
PR	17-NOV-1998;	98US-0108806.
PR	17-NOV-1998;	98US-0108807.
PR	17-NOV-1998;	98US-0108867.
PR	17-NOV-1998;	98US-0108925.
PR	18-NOV-1998;	98US-0108848.
PR	18-NOV-1998;	98US-0108849.
PR	18-NOV-1998;	98US-0108850.
PR	18-NOV-1998;	98US-0108851.
PR	18-NOV-1998;	98US-0108852.
PR	18-NOV-1998;	98US-0108858.
PR	18-NOV-1998;	98US-0108904.
XX		
XX	(GETH) GENENTECH INC.	
XX		
PI	Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;	
XX	WPI, 2000-237871/20.	
XX	P-PSDB; AA939380.	
OR		

XX	New mammalian DNA sequences encoding transmembrane, receptor or
PT	secreted PRO polypeptides, useful for screening of potential peptide or
XX	small molecule inhibitors of the relevant receptor/ligand interactions
PS	
XX	Claim 2; Fig 81; 773pp; English.
CC	AAA37022 to AAA37144 encode the new isolated human transmembrane,
CC	receptor or secreted PRO polypeptides given in AA993440 to AA99462. The
CC	transmembrane and receptor PRO proteins can be used for screening of
CC	potential peptide or small molecule inhibitors of the relevant
CC	receptor/ligand interactions. The polypeptides and nucleotide sequences
CC	encoding then have various industrial applications, including uses as
CC	pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
CC	PCR primers and hybridisation probes used in the isolation of the PRO
CC	polypeptides from the present invention.
XX	
SD	Sequence 771 BP; 169 A; 231 C; 195 G; 176 T; 0 other:
	Alignment Scores:
	Pred. No.: 5,68e-111 Length: 771
	Score: 117.00 Matches: 117
	Percent Similarity: 100.00% Conservative: 0
	Best Local Similarity: 100.00% Mismatches: 0
	Query Match: 100.00% Indels: 0
	DB: 21 Gaps: 0
US-09-092-297-17 (1-117) x AAA37062 (1-771)	
QY	1 ProLeuGlnProProAlaGAlaMeAlaProArgGlyCysIleValAlaAlaPheAlaIle 20
Db	3 CCACGCAACACCCAGAGCCATGGCTCCCGAGGCGCATCGTAGCTGCTTGGCAATT 62
QY	21 PheCysIleSerArgLeuLeuCysSerHisGlyValaProValaAlaProMetThrProTyr 40
Db	63 TTTCGATCTCCAGAGCTCTCTGCTCACAGGAGCCCCCAAGTGGCCCCCAAGACTCCCTTAC 122
QY	41 LeuMetLeuCysGlnProHisIlyAsrCysGlyAspIlySerPheTyrAspProLeuGlnHis 60
Db	123 CTGATGCTGCTGCCAGCCACACAAAGATGTGGGACAAAGTTCTACAGACCCCTGCAGCAC 182
QY	61 CysCysTyrAspAspAlaValaValaProLeuAlaArgThrGlnThrCysGlyAsnCysThr 80
Db	183 TGTTCATATGATGATGCCGCTGCTGCCCTTGGCCAGGACCCAGACGATGGAAATCGCACCC 242
QY	81 PheArgValCysPheGlnGlnCysCysProTyrThrPheMetAlaIlyIleuIleAsnGln 100
Db	243 TTCAGAGCTGCTTGTAGACAGTGTGCCCCCTGAGACCTTCACTGTGAAGTGATAAACACAG 302
QY	101 AsnCysAspSerAlaArgThrSerAspAspArgLeuCysArgSerValSer 117
Db	303 AACGTCAGCTCAGCCCGGACCTCGGATGACAGAGCTTTGTCGCGAGTGTACGC 353
RESULT 6	
AAAS21491	
AC	AAAS21491 standard; cDNA; 771 BP.
XX	
XX	AAAS21491;
XX	
XX	24-OCT-2001 (first entry)
DE	Human cDNA sequence encoding for PRO1274 polypeptide.
KW	Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW	breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW	cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW	adipocyte; A-peptide; factor VIIA; gene therapy; ss.
XX	
OS	Homo sapiens.
XX	
XX	WO200140466-A2.
PN	
XX	07-JUN-2001.
DD	

XX 01-DEC-2000; 2000MO-US32678.
 PF 01-DEC-1999; 99MO-US28301.
 XX 01-DEC-1999; 99MO-US28634.
 PR 02-DEC-1999; 99MO-US28514.
 PR 02-DEC-1999; 99MO-US28564.
 PR 02-DEC-1999; 99MO-US28565.
 PR 09-DEC-1999; 99US-0170262.
 PR 16-DEC-1999; 99MO-US30095.
 PR 20-DEC-1999; 99MO-US30911.
 PR 20-DEC-1999; 99MO-US30999.
 PR 30-DEC-1999; 99MO-US31243.
 PR 06-JAN-2000; 2000MO-US00277.
 PR 06-JAN-2000; 2000MO-US00376.
 PR 11-FEB-2000; 2000MO-US03565.
 PR 18-FEB-2000; 2000MO-US04341.
 PR 18-FEB-2000; 2000MO-US04342.
 PR 22-FEB-2000; 2000MO-US04414.
 PR 24-FEB-2000; 2000MO-US04914.
 PR 24-FEB-2000; 2000MO-US05004.
 PR 01-MAR-2000; 2000MO-US05601.
 PR 20-MAR-2000; 2000MO-US07377.
 PR 21-MAR-2000; 2000MO-US07532.
 PR 30-MAR-2000; 2000MO-US08439.
 PR 17-MAY-2000; 2000MO-US13705.
 PR 22-MAY-2000; 2000MO-US14042.
 PR 30-MAY-2000; 2000MO-US14941.
 PR 02-JUN-2000; 2000MO-US15264.
 PR 10-NOV-2000; 2000MO-US30873.
 XX (GETH) GENENTECH INC.
 PA Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerlitsen ME, Goddard A, Godowski PJ, Gurney AL, Sherrwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 DR WPI: 2001-408281/43.
 DR P-PSDB; AAU12419.
 XX Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PT lung, breast, prostate, cervical.
 XX Claim 3: Fig 495; 813pp; English.
 PS AAS21244-AAS21518 encode for novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 XX Sequence 771 BP; 169 A; 231 C; 195 G; 176 T; 0 other;
 Alignment Scores: 5.68e-111 Length: 771
 Pred. No.: 117.00 Matches: 117
 Score:

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-092-297-17 (1-117) x AAS21491 (1-771)
 QY 1 ProLeuGlnProProArgAlaMetAlaProArgGlyCysIleValAlaValPheAlaIle 20
 DB 3 CCACTGCACACACCCAGAGCCATGCTCCCGAGGCTCATGTAGTGTGGCCATT 62
 QY 21 PheCysIleSerArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
 DB 63 TTCTGCATCTCCAGGCTCTCTGCTGCACACAGGAGCCAGGCGCCCATACCTTAC 122
 QY 41 LeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
 DB 123 CTGATGCTGTGCCAGCCACACAGAGATGTGGGACAAAGTCTACGACCCCTGCAGCAC 182
 QY 61 CysCysTyrAspAspAlaValAlaValProLeuAlaArgThrInThrCysGlyAsnCysThr 80
 DB 183 TGTTCGTAATGATGATGCCGTGCTGCCAGGACCCAGAGCTGTGAAACTGCACC 242
 QY 81 PheArgValCysPheGluGlnCysCysProTrpThrPheMetValIleLeuIleAsnGln 100
 DB 243 TTCAGAGTCGCTTGAAGCAGTGTGCCCTCGACCTTCATGCTGAACCTGATAACACAG 302
 QY 101 AsnCysAspSerAlaArgThrSerAspAspArgLeuLeuLeuLeuLeuLeuLeu 117
 DB 303 AACTGCAGCTCAGCCGAGACCTCGAGTACAGGCTTGTGTCCAGTGTCCAG 353
 RESULT 7
 AAC85962
 ID AAC85962 standard; cDNA; 771 BP.
 XX
 AC AAC85962;
 XX
 DT 22-AUG-2001 (first entry)
 XX
 DE Native sequence of PRO1274 cDNA, clone DNM64889-1541.
 XX
 KW PRO: PRO1081; PRO1274; PRO10272; proliferation; T-lymphocyte;
 KW PRO1199; PRO1556; PRO4401; PRO10268; inhibition; stimulation;
 KW infiltration; mononuclear cell; eosinophils; erythema multiforme;
 KW polymorphonuclear neutrophils; PMN; antibody; immune-related disorder;
 KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
 KW juvenile chronic arthritis; spondyloarthritis; systemic sclerosis;
 KW idiopathic inflammatory myopathy; Sjogren's syndrome; skin disease;
 KW autoimmune vasculitis; sarcoidosis; autoimmune haemolytic anaemia; asthma;
 KW autoimmune thrombocytopaenia; thyroiditis; diabetes mellitus; allergy;
 KW immune-mediated renal disease; demyelination; central nervous system;
 KW peripheral nervous system; idiopathic demyelinating polyneuropathy;
 KW Guillain-Barre syndrome; hepatobiliary disease; eosinophilic pneumonia;
 KW chronic active hepatitis; primary biliary cirrhosis; allergic rhinitis;
 KW granulomatous hepatitis; sclerosing cholangitis; food hypersensitivity;
 KW inflammatory bowel disease; gluten-sensitive enteropathy; urticaria;
 KW Whipple's disease; idiopathic pulmonary fibrosis; contact dermatitis;
 KW psoriasis; atopic dermatitis; hypersensitivity pneumonitis;
 KW graft rejection; graft-versus-host disease; ss.
 KW
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 24..356
 FT /tag- a
 FT /product- "PRO1274 polypeptide"
 FT 24..95
 FT /tag- b
 FT mat_peptide 96..353
 FT /tag- c
 XX
 PN WO200104065-A2
 XX

PD 07-JUN-2001.
XX
PF 10-NOV-2000; 2000MO-US30873.
XX
PR 30-NOV-1999; 99MO-US28313.
PR 09-DEC-1999; 99US-0170262.
PR 23-DEC-1999; 99US-0172059.
PR 11-JAN-2000; 2000US-0175481.
PR 20-JAN-2000; 2000US-0177118.
PR 18-FEB-2000; 2000MO-US04342.
PR 03-MAR-2000; 2000US-0187202.
PR 30-MAY-2000; 2000MO-US14941.
PR 05-JUN-2000; 2000US-0209832.
PR 24-AUG-2000; 2000MO-US23328.
XX
PA (GETH) GENENTECH INC.
PI Fong S, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ,
PI Tumas D, Watanabe CK, Wood WI, Zhang Z;
PI
PI WPI: 2001-381384/40.
DR P-PSDB; AAB47292.
XX
XX Isolated PRO polypeptide useful for treat or diagnose an immune-related
XX disorder e.g. arthritis, asthma, allergy, diabetes or psoriasis -
XX
XX Claim 2; Fig 3; 14pp; English.
XX
XX The sequences given in AAC85961-69 encode PRO polypeptides. PRO1081,
CC PRO1274 and PRO10272 stimulate the proliferation of T-lymphocytes and
CC PRO1199, PRO1556, PRO4401 and PRO10268 inhibit the proliferation of
CC T-lymphocytes. PRO1754 and PRO9912 act to enhance the infiltration of
CC mononuclear cells, eosinophils or polymorphonuclear neutrophils (PMN)
CC into the tissue of a mammal. The PRO cDNA's and antibodies which
CC bind to them, are used to treat an immune-related disorder in a
CC mammal. Such disorders include systemic lupus erythematosus,
CC rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis, a
CC myopathy, Sjogren's syndrome, systemic sclerosis, an idiopathic inflammatory
CC autoimmune haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis,
CC diabetes mellitus, immune-mediated renal disease, a demyelinating
CC disease of the central or peripheral nervous system, idiopathic
CC demyelinating polyneuropathy, Guillain-Barre syndrome, a chronic
CC inflammatory demyelinating polyneuropathy, a hepatobiliary disease,
CC infectious or autoimmune chronic active hepatitis, primary biliary
CC cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bulous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an
CC hypersensitivity, asthma, allergic rhinitis, atopic dermatitis, food
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease.
XX
XX Sequence 771 BP; 169 A; 231 C; 195 G; 176 T; 0 other;
XX
XX Alignment Scores:
Pred. No.: 5.68e-111 Length: 771
Score: 117.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0
US-09-092-297-17 (1-117) x AAC85962 (1-771)
QY 1 ProlaughinProArgAlaMetLarProArgGlyCysIleValAlaValAlaIle 20
DB 3 CCACGCAACCAACCAAGCAGCATGCTCCCGAGGCTGCATGCTCTTGGCATT 62
QY 21 PheCysIleSerArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
|||||

DB 63 TTCGCACTCCAGGCTCTCTGCTACACAGAGCCCGCCAGTGGCCCCCATGATCTCTTAC 122
QY 41 LeuMetLeuCysGlnProHisLysArgCysGlyAspLysPheTyrAspProLeuGlnHis 60
DB 123 CTGATGCTGTGCGACGCCACACAGAGATGTGGGACAACTTCTACAGCCCTTCAGCAC 182
QY 61 CysCysTyrAspAspAlaValAlaProLeuAlaArgTyrGlnTyrCysGlyAsnCysThr 80
DB 183 TGTTCCTATGATGATGATGCGCTGCTGCTGGCCAGAGACCAGAGCGTGTGGAACCTGACAC 242
QY 81 PheArgValCysPheGlnGlnCysCysProTrrpThrPheMetValLysLeuIleAsnGln 100
DB 243 TTCAGAGTCTCTTGTACAGCTGCTGCCCTCGAGCCTTCATGATGTAAGCTGATAACACAG 302
QY 101 AsnCysAspSerAlaArgThrSerAspAspArgLeuLeuLeuLeuLeuLeuLeu 117
DB 303 AACGCGACTGACGCCCGGACCTCGGATGACAGGCTTGTCCAGTGTACG 353
RESULT 8
AAF54298
ID AAF54298 standard; DNA; 771 BP.
XX
XX AAF54298;
AC
XX
XX 02-APR-2001 (first entry)
DT
XX
XX DNA encoding protein of the invention #41.
DE
XX
XX Secreted; transmembrane; gene therapy; ss.
KW
XX
XX Unidentified.
OS
XX
XX WO200078961-A1.
PN
XX
XX 28-DEC-2000.
PD
XX
XX 18-FEB-2000; 2000MO-US04342.
PF
XX
XX 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 01-SEP-1999; 99MO-US20111.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99MO-US28313.
PR 02-DEC-1999; 99MO-US28551.
PR 16-DEC-1999; 99MO-US30095.
PR 05-JAN-2000; 2000MO-US00219.
PR 06-JAN-2000; 2000MO-US00376.
XX
XX (GETH) GENENTECH INC.
PA
XX
XX Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S,
PI Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ,
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D,
PI Watanabe CK, Williams PW, Wood WI;
PI
PI WPI: 2001-071395/08.
DR
XX
XX Secreted and transmembrane proteins and nucleic acids designated PRO,
PT useful as hybridization probes, in chromosome and gene mapping and gene
PT therapy -
XX
XX Claim 2; Fig 81; 787pp; English.
XX
XX The present invention relates to secreted and transmembrane proteins.
CC These proteins and the DNA encoding them may be used as hybridization
CC probes, in chromosome and gene mapping and in the generation of
CC anti-sense RNA and DNA. They may also be used to generate either
CC transgenic animals or knockout animals which are in turn useful for
CC development and screening of therapeutically useful reagents.
CC The nucleic acids may also be used in gene therapy.
XX
XX Sequence 771 BP; 169 A; 231 C; 195 G; 176 T; 0 other;
SQ

PD 06-JAN-2000.
 XX
 XX 25-JUN-1999; 99MO-US14484.
 XX
 XX 26-JUN-1998; 98US-0090762.
 PR 31-JUL-1998; 98US-0094983.
 PR 01-OCT-1998; 98US-0102686.
 PR 11-DEC-1998; 98US-0112129.
 XX
 PA (INCY-) INCYTE PHARM INC.
 PI Lal P, Tang YF, Gorgone GA, Corley NC, Guegler KJ, Baughn MR,
 PI Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL,
 PI Bandman O;
 DR WPI; 2000-160673/14.
 DR P-PSDB; AAY87273.
 XX
 XX New human signal peptide-containing proteins useful in treatment,
 PT prevention and diagnosis of e.g. cancer, inflammation and
 PT cardiovascular disease
 PT
 PS Claim 9; Page 281; 327pp; English.
 CC AA298109 to AA298242 encode AAY87224 to AAY87357 which represent the
 CC human signal peptide-containing proteins HSP-1 to HSP-134. HSPs have
 CC anticancer, anti-inflammatory, antimicrobial, neurotrophic, hepatotropic,
 CC neuroprotective, cardiovascular and antisthmatic activities, and can
 CC be used in gene therapy. HSPs can be used to treat or prevent disorders
 CC associated with decreased activity or function of HSP. Antagonists of
 CC HSP are used to treat or prevent disorders associated with increased
 CC activity or function of HSP. Such diseases include cell proliferation
 CC (including cancer), inflammation, cardiovascular, neurological,
 CC reproductive or developmental disorders, (e.g. arteriosclerosis,
 CC cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,
 CC asthma, Crohn's disease, microbial or other infections, congestive or
 CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's
 CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSP
 CC nucleic acids can be used for the recombinant production of HSP, for
 CC detecting HSP in standard hybridisation and amplification assays (for
 CC diagnosis and monitoring), in gene therapy, as antisense,
 CC triplex-forming or ribozyme therapeutics, for detecting related sequences
 CC or genetic variations, and for chromosomal mapping. HSP are also used to
 CC raise specific antibodies (Ab) and to screen for agonists and
 CC antagonists (potential therapeutic agents). Ab are used to diagnose, or
 CC monitor, HSP-related diseases (in usual immunoassays), as therapeutic
 CC antagonists, in competitive drug screens, and for purification of HSP
 CC from natural sources.
 XX
 XX Sequence 748 BP; 159 A; 227 C; 193 G; 169 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 5,88e-110 Length: 748
 Score: 116.00 Matches: 116
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.15% Indels: 0
 DB: 21 Gaps: 0
 US-09-092-297-17 (1-117) x AA298158 (1-748)
 OY 2 LeuGlnProProIaValAlaMetAlaProArgGlyCysIleValAlaValAlaPhe 21
 DB 1 CTGCACACACCCAGAGCCATGGCTCCCGAGCGTCATGCTGCTTTGGCAATTTC 60
 OY 22 CysIleSerArgLeuLeuLeuSerHisGlyAlaProValAlaProMetThrProTyrLeu 41
 DB 61 TGCATCTCCAGGCTCTCTGCTACACAGGAGCCAGTGGCCCATGACTCTTACCTG 120
 OY 42 MetLeuGlyGlnProHisLysArgCysGlyAspLysPheTyrAspProLeuGlnHisCys 61
 DB 121 ATGCTGTGCAGACGACACAGAGATGTGGGAGACAGTCTACGACCCCTGACGACTGT 180

OY 62 CysTyrAspAspAlaValAlaProLeuAlaArgThrGlnThrCysGlyAsnCysThrPhe 81
 DB 181 TGTATGATGATGCCGTGCTGCCCTTGCCAGAGACCCAGAGCTGTGAACTGCACCTTC 240
 OY 82 ArgValCysPheGluGlnCysCysProThrPheMetValLysLeuIleAsnGlnAsn 101
 DB 241 AGAGTCTGCTTTGAGCAGTGTCTGCCCTGAGACCTTCATGTGAAGCTGATTAACAGAC 300
 OY 102 CysAspSerAlaArgThrSerAspArgLeuGlyAsnArgSerValSer 117
 DB 301 TGCACCTCAGCCCGAGACTCGATGACAGCTTGTGTCAGATGTGAC 348
 RESULT 11:
 ID AAX01683 standard; DNA; 196 BP.
 AC AAX01683;
 XX
 XX 21-APR-1999 (first entry)
 DE Urinary tract tissue library BL172 gene specific cDNA clone 1554838.
 XX
 XX BL172; urinary tract; marker; cancer; recombinant; human; ss.
 KM
 XX Homo sapiens.
 OS
 XX
 XX W09855656-A1.
 PN
 PD 10-DEC-1998.
 XX
 XX 05-JUN-1998; 98MO-US11693.
 PF
 XX
 XX 05-JUN-1997; 97US-0869579.
 PR
 XX (ABBO) ABBOTT LAB.
 PA
 XX Billing-medel PA, Cohen M, Colpitts TL, Friedman PN,
 PI Gordon J, Granadosen, Hodges SC, Klass MR, Kratochvil JD;
 PI Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;
 DR WPI; 1999-045802/04.
 XX
 XX New purified polynucleotide BL172 derivatives and encoded
 PT polypeptides - useful in the identification of markers, which are
 PT indicative of urinary tract diseases or conditions
 PT
 XX Claim 1; Page 89; 114pp; English.
 PS
 XX Sequences AAX01683-85 represent overlapping clones specific for urinary
 CC tract tissue library BL172. The BL172 polynucleotides are used to detect
 CC target BL172 polynucleotides and BL172 mRNA and the BL172 polypeptides
 CC are used to detect BL172 antigens/antibodies in a test sample. The
 CC identification of certain markers in these methods are indicative of the
 CC presence of urinary tract disease, especially cancer. Additionally, the
 CC polypeptides are used to detect antibodies that bind specifically to a
 CC BL172 epitope derived from a BL172 polypeptide. Host cells containing a
 CC recombinant expression vector comprising the BL172 polynucleotide
 CC sequences are used to produce the polypeptides containing BL172
 CC epitopes. The methods aid in diagnosis, detection, staging, monitoring,
 CC prognostication, in vivo imaging, prevention and treatment of diseases
 CC or conditions associated with BL172, especially urinary tract cancer.
 XX
 XX Sequence 196 BP; 38 A; 70 C; 44 G; 44 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 4.34e-57 Length: 196
 Score: 64.00 Matches: 64
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 54.70% Indels: 0
 DB: 20 Gaps: 0
 US-09-092-297-17 (1-117) x AAX01683 (1-196).

QY 1 ProLeuGlnProProArgAlaMetAlaProArgGlyCysIleValAlaValPheAlaIle 20
 DB 3 CCAGTGCACACCCAGACCCAGTGGCTCCCGAGGCTGCATGCTGCTTGGCATT 62
 QY 21 PheCysIleSerArgLeuLeuGlnCysSerHisGlyAlaProValAlaProMetThrProTyr 40
 DB 63 TTCTGCATCTCCAGGCTCTTCTGCTCAGGAGGCCCGGAGCCCGCCCTGACTCTTAC 122
 QY 41 LeuMetLeuGlnProHisLysArgCysGlyAspLysPheTyrAspProLeuGlnHis 60
 DB 123 CCGAGGCTGTGCAGCCACACAGAGATGTGGGACAACTTCTAGACCCCTGCAGCAC 182
 QY 61 CysCysTyrAsp 64
 DB 183 TGTGTCTATGAT 194
 RESULT 12
 AAX01684
 ID AAX01684 standard; DNA: 236 BP.
 AAX01684;
 21-APR-1999 (first entry)
 DE Urinary tract tissue library BL172 gene specific cDNA clone 2624118.
 XX BL172; urinary tract; marker; cancer; recombinant; human; ss.
 KM Homo sapiens.
 XX W09855656-A1.
 PN 10-DEC-1998.
 PD 05-JUN-1998; 98WO-US11693.
 PF 05-JUN-1997; 97US-0869579.
 PR 05-JUN-1997; 97US-0869579.
 XX (ABBO) ABBOTT LAB.
 PA Billing-medel PA, Cohen M, Colpitts TL, Friedman PN;
 PI Gordon J, Grandosen, Hodges SC, Klass MR, Katochwal JD;
 PI Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;
 XX WPI: 1999-045802/04.
 DR New purified polynucleotide BL172 derivatives and encoded
 PT polypeptides - useful in the identification of markers, which are
 PT indicative of urinary tract diseases or conditions
 XX Claim 1; Page 89; 114pp; English.
 PS Sequences AAX01683-85 represent overlapping clones specific for urinary
 CC tract tissue library BL172. The BL172 polynucleotides are used to detect
 CC target BL172 polynucleotides and BL172 mRNA and the BL172 polypeptides
 CC are used to detect BL172 antigens/antibodies in a test sample. The
 CC identification of certain markers in these methods are indicative of the
 CC presence of urinary tract disease, especially cancer. Additionally, the
 CC polypeptides are used to detect antibodies that bind specifically to a
 CC BL172 epitope derived from a BL172 polypeptide. Host cells containing a
 CC recombinant expression vector comprising the BL172 polynucleotide
 CC sequences are used to produce the polypeptides containing BL172
 CC epitopes. The methods aid in diagnosis, detection, staging, monitoring,
 CC prognostication, in vivo imaging, prevention and treatment of diseases
 CC or conditions associated with BL172, especially urinary tract cancer.
 XX Sequence 236 BP; 49 A; 69 C; 62 G; 56 T; 0 other;
 SO Alignment Scores:
 Pred. No.: 5,17e-57 Length: 236
 Score: 64.00 Matches: 64
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 54.70% Indels: 0
 DB: 20 Gaps: 0
 US-09-092-297-17 (1-117) x AAX01684 (1-236)
 QY 54 PheTyrAspProLeuGlnHisCysCysTyrAspAspAlaValAlaProLeuAlaArgThr 73
 DB 1 TTCTAGACACCCCTGCAGACACTGTGCTATGATGATGCCGTGCTGCTGCAGGACC 60
 QY 74 GlnThrCysGlyAsnCysThrPheArgValCysPheGlnGlnCysCysProThrPhe 93
 DB 61 CAGAGGTGGAGAACTGCACCTTCAGAGCTCTTGGACCACTGCTGCCCTGGACCTTC 120
 QY 94 MetValLysLeuIleAsnGlnAsnCysAspSerAlaArgThrSerAspAspArgLeuGln 113
 DB 121 ATGCTGAGACGTATTAACACAGACTGCCACTGACCCCGGACCTCGAGATGACAGCTTGT 180
 QY 114 ArgSerValSer 117
 DB 181 CGCACTGTCAGC 192
 RESULT 13
 AAS92056
 ID AAS92056 standard; cDNA: 1338 BP.
 AAS92056;
 13-FEB-2002 (first entry)
 DE DNA encoding novel human diagnostic protein #27860.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX Homo sapiens.
 PN W0200175067-A2.
 PD 11-OCT-2001.
 PF 30-MAR-2001; 2001WO-US08631.
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 PI P-PSDB: ABG27869.
 DR WPI: 2001-639362/73.
 DR P-PSDB: ABG27869.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX Claim 1; SEQ ID NO 27860; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. A664197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1338 BP; 311 A; 350 C; 348 G; 329 T; 0 other;
Alignment Scores:
Pred. No.: 8.5 Length: 1338
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.69% Indels: 0
DB: Gaps: 0
US-09-092-297-17 (1-117) x AAS92056 (1-1338)
QY 78 AaNCysThrPheArgValCysPheGlu 86
DB 1266 AACTGCACCTTCAGAGTCGCTCGAG 1292
RESULT 14
ABA49458
ID ABA49458 standard; DNA; 252 BP.
XX
AC ABA49458;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #8153.
XX
KW Human; microarray; single exon probe; gene expression; breast;
XX disease; cancer; ss.
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
XX useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes
XX
PS Claim 4; SEQ ID NO 8153; 327bp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to

CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 252 BP; 82 A; 53 C; 54 G; 63 T; 0 other;
Alignment Scores:
Pred. No.: 18.6 Length: 252
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.84% Indels: 0
DB: Gaps: 0
US-09-092-297-17 (1-117) x ABA49458 (1-252)
QY 70 LeuAlaArgThrGlnThrCysGly 77
DB 112 CTGGCTAGAACCCAGACTTGTGGC 135
RESULT 15
ABA67371
ID ABA67371 standard; DNA; 252 BP.
XX
AC ABA67371;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #15676.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human foetal liver
XX
PS Claim 4; SEQ ID NO 15676; 639bp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at http://wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 252 BP; 82 A; 53 C; 54 G; 63 T; 0 other;

Alignment Scores:

Pred. No.:	18.6	Length:	252
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.84%	Indels:	0
DB:	22	Gaps:	0

US-09-092-297-17 (1-117) x ABA67371 (1-252)

QY 70 LeuAlaArgThrGlnThrCysGly 77
|||||
DB 112 CTGCGTAGAACCCAGACTGTGTGGC 135

Search completed: November 3, 2002, 05:05:08
Job time : 215 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 3, 2002, 03:21:26 ; Search time 46 Seconds

(without alignments)
624.763 Million cell updates/sec

Title: US-09-092-297-17

Perfect score: 117

Sequence: 1 PLUOPPRAMPKRCIVAVFAI.....INONCDARTSDRLCRSVS 117

Scoring table:

	OLIGO
Xgapop 60.0	Xgapext 60.0
Ygapop 60.0	Ygapext 60.0
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 383533 seqs, 122816752 residues

Word size: 1

Total number of hits satisfying chosen parameters: 763587

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=x1h
-O=/cg2n2_1/USPTO.spool/US09092297/runat.29102002.094017.22411/app_query.fasta.1.263
-DB=Issued_Patents_NA -QFWT=fastap -SUFFIX=olip2n.rn1 -MINMATCH=0.1 -DOOPCL=0
-DOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09092297 -CGCN1.1.13 @runat.29102002.094017.22411 -NCPUS=6 -ICPU=3
-NO_XLPXY -NO_KMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database: Issued_Patents_NA.*

1: /cg2n2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cg2n2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cg2n2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cg2n2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cg2n2_6/ptodata/1/lna/PCITUS.COMB.seq:*
6: /cg2n2_6/ptodata/1/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	6.0	846	4	US-08-998-416-325
2	7	6.0	1753	4	US-08-205-697A-6
3	7	6.0	1753	4	US-08-702-5525-6
4	7	6.0	1753	5	PCT-US95-02576-6
5	7	6.0	2351	2	US-08-548-159-6
6	7	6.0	2856	4	US-08-716-873-3
7	7	6.0	2856	4	US-08-716-873-13
8	7	6.0	2856	4	US-08-716-873-14
9	7	6.0	2856	4	US-09-368-431-3
10	7	6.0	2856	4	US-09-368-431-13
11	7	6.0	2856	4	US-09-368-431-14
12	7	6.0	3715	4	US-09-041-886-10

C	13	7	6.0	4014	4	US-09-541-782-1	Sequence 1, App1
C	14	7	6.0	4080	1	US-08-570-311-7	Sequence 7, App1
C	15	7	6.0	4080	2	US-08-353-485-7	Sequence 7, App1
C	16	7	6.0	4140	4	US-08-716-873-1	Sequence 1, App1
C	17	7	6.0	4140	4	US-09-368-431-1	Sequence 1, App1
C	18	7	6.0	6241	1	US-08-570-311-25	Sequence 25, App1
C	19	7	6.0	7266	3	US-08-336-308A-9	Sequence 9, App1
C	20	7	6.0	7266	3	US-08-832-324-5	Sequence 5, App1
C	21	7	6.0	7266	4	US-09-480-831-9	Sequence 9, App1
C	22	7	6.0	8640	1	US-08-570-311-28	Sequence 28, App1
C	23	7	6.0	8982	3	US-08-976-255-5	Sequence 5, App1
C	24	7	6.0	11558	5	PCT-US93-06251-23	Sequence 23, App1
C	25	7	6.0	26664	4	US-09-564-805-28	Sequence 28, App1
C	26	7	6.0	32207	2	US-08-770-379-20	Sequence 20, App1
C	27	7	6.0	32207	4	US-08-757-669A-20	Sequence 20, App1
C	28	7	6.0	32207	4	US-09-230-371A-20	Sequence 20, App1
C	29	7	6.0	32207	4	US-09-488-856A-61	Sequence 61, App1
C	30	6	5.1	37	2	US-08-468-700-17	Sequence 17, App1
C	31	6	5.1	37	1	US-08-468-700-17	Sequence 17, App1
C	32	6	5.1	37	2	US-08-704-706A-17	Sequence 17, App1
C	33	6	5.1	37	2	US-08-985-659-17	Sequence 17, App1
C	34	6	5.1	40	4	US-09-275-850-154	Sequence 154, App1
C	35	6	5.1	41	1	US-07-931-473B-280	Sequence 280, App1
C	36	6	5.1	41	1	US-07-714-131C-280	Sequence 280, App1
C	37	6	5.1	41	1	US-08-412-110-280	Sequence 280, App1
C	38	6	5.1	41	1	US-08-409-442A-280	Sequence 280, App1
C	39	6	5.1	41	2	US-08-469-609A-280	Sequence 280, App1
C	40	6	5.1	41	3	US-09-143-190-280	Sequence 280, App1
C	41	6	5.1	43	1	US-07-931-473B-306	Sequence 306, App1
C	42	6	5.1	43	1	US-07-714-131C-306	Sequence 306, App1
C	43	6	5.1	43	1	US-08-412-110-306	Sequence 306, App1
C	44	6	5.1	43	1	US-08-409-442A-306	Sequence 306, App1
C	45	6	5.1	43	2	US-08-469-609A-306	Sequence 306, App1

ALIGNMENTS

RESULT 1
US-08-998-416-325
Sequence 325, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESSES:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 325:
SEQUENCE CHARACTERISTICS:
LENGTH: 846 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1060UP
US-08-998-416-325

Alignment Scores:
Pred. No.: 126 Length: 846
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 4 Gaps: 0

US-09-092-297-17 (1-117) x US-08-998-416-325 (1-846)

QY 65 AspaIaValaIproleuAa 71
DB 508 GACGGGCTGCTGCTCCAC 528

RESULT 2
US-08-205-697A-6
Sequence 6, Application US/08205697A
Patent No. 6218510
GENERAL INFORMATION:
APPLICANT: Sharpe, Arlene H.
APPLICANT: Bortello, Francescopaulo
APPLICANT: Freeman, Gordon J.
APPLICANT: Nadler, Lee M.
TITLE OF INVENTION: NO. 6218510el Forms of T Cell Costimulatory Molecules
TITLE OF INVENTION: and Uses Therefor
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHYE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/205,697A
FILING DATE: 02-Mar-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: BWI-120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1753 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-205-697A-6

Alignment Scores:
Pred. No.: 250 Length: 1753
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 4 Gaps: 0

US-09-092-297-17 (1-117) x US-08-205-697A-6 (1-1753)

QY 24 SerArgIeuCysSerHis 30
DB 274 TCCGACTCTCTGCTCCAC 294

RESULT 3
US-08-702-525-6
Sequence 6, Application US/08702525
Patent No. 6294660
GENERAL INFORMATION:
APPLICANT: Sharpe, Arlene H.
APPLICANT: Bortello, Francescopaulo
APPLICANT: Freeman, Gordon
APPLICANT: Nadler, Lee
TITLE OF INVENTION: NO. 6294660el Forms of T Cell Costimulatory
TITLE OF INVENTION: Molecules and Uses Therefor
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHYE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,525
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/205,697
FILING DATE: 02-Mar-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: BWI-120CUPS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1753 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-702-525-6

Alignment Scores:
Pred. No.: 250 Length: 1753
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 4 Gaps: 0

US-09-092-297-17 (1-117) x US-08-702-525-6 (1-1753)

QY 24 SerArgIeuCysSerHis 30
DB 274 TCCGACTCTCTGCTCCAC 294

RESULT 4
PCT-US95-02576-6
Sequence 6, Application PC/TUS9502576
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Novel Forms of T Cell Costimulatory Molecules
TITLE OF INVENTION: And Uses Therefor
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02576
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/205,697
FILING DATE: 02-Mar-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: BWI-120CPCC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1753 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
PCT-US95-02576-6

Alignment Scores:
Pred. No.: 250 Length: 1753
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: Gaps: 0

US-09-092-297-17 (1-117) x PCT-US95-02576-6 (1-1753)

OY 24 SerArgLeuCySerHis 30
|||||
Db 274 TCCGACCTCTGCTCCAC 294

RESULT 5
US-08-548-159-6/c
Sequence 6, Application US/08548159
Patent No. 5989551
GENERAL INFORMATION:
APPLICANT: MacLaren, No. 5989551 K.
APPLICANT: No. 5989551 Kins, Abner L.
APPLICANT: Lan, Michael S.
APPLICANT: Li, Qing
TITLE OF INVENTION: Materials and Methods for Detection and
TITLE OF INVENTION: Treatment of Insulin Dependent Diabetes
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville

STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/548,159
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Bencen, Gerard H.
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: UP154.C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2351 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Mouse IA-2 Beta cDNA
US-08-548-159-6

Alignment Scores:
Pred. No.: 330 Length: 2351
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: Gaps: 0

US-09-092-297-17 (1-117) x US-08-548-159-6 (1-2351)

OY 111 ArgLeuCyArgSerValSer 117
|||||
Db 80 AGACTCTGCAGGAGTGTCTCA 60

RESULT 6
US-08-716-873-3
Sequence 3, Application US/08716873
Patent No. 6194166
GENERAL INFORMATION:
APPLICANT: Takashi OKADO et al.
TITLE OF INVENTION: GENE REGULATING AUREOBASIDIN SENSITIVITY
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/716,873
FILING DATE: September 20, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.

REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2856 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-716-873-3

Alignment Scores:
Pred. No.: 396 Length: 2856
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 4 Gaps: 0

US-09-092-297-17 (1-117) x US-08-716-873-3 (1-2856)

QY 23 IlleSeraGlauLeucCysser 29
Db 1224 ATCAGTCGCGCTGTGTGTCG 1244

RESULT 7
US-08-716-873-13/C
Sequence 13, Application US/08716873
Patent No. 6194166
GENERAL INFORMATION:
APPLICANT: Takashi OKADO et al.
TITLE OF INVENTION: GENE REGULATING AUREOBASIDIN SENSITIVITY
NUMBER OF SEQUENCES: 50
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/716,873
FILING DATE: September 20, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 2856 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ANTI-SENSE: Yes
US-08-716-873-13

Alignment Scores:
Pred. No.: 396 Length: 2856
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 4 Gaps: 0

US-09-092-297-17 (1-117) x US-08-716-873-13 (1-2856)

QY 23 IlleSeraGlauLeucCysser 29
Db 1633 ATCAGTCGCGCTGTGTGTCG 1613

RESULT 8
US-08-716-873-14/C
Sequence 14, Application US/08716873
Patent No. 6194166
GENERAL INFORMATION:
APPLICANT: Takashi OKADO et al.
TITLE OF INVENTION: GENE REGULATING AUREOBASIDIN SENSITIVITY
NUMBER OF SEQUENCES: 50
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/716,873
FILING DATE: September 20, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2856 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: mRNA
ANTI-SENSE: Yes
US-08-716-873-14

Alignment Scores:
Pred. No.: 396 Length: 2856
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 4 Gaps: 0

US-09-092-297-17 (1-117) x US-08-716-873-14 (1-2856)

QY 23 IlleSeraGlauLeucCysser 29
Db 1633 ATCAGTCGCGCTGTGTGTCG 1613

RESULT 9
US-09-368-431-3
Sequence 3, Application US/09368431
Patent No. 6294651
GENERAL INFORMATION:
APPLICANT: Takashi OKADO et al.
TITLE OF INVENTION: GENE REGULATING AUREOBASIDIN SENSITIVITY
TITLE OF INVENTION: (AS AMENDED)
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: WENDROTH, LIND & PONACK, L.L.P.
STREET: 2033 K Street, N.W., #800
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/368,431
FILING DATE: August 5, 1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/716,873
FILING DATE: September 20, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2856 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-368-431-3
Alignment Scores:
Pred. No.: 396
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 5.98%
DB: 4
Gaps: 0
US-09-092-297-17 (1-117) x US-09-368-431-3 (1-2856)
QY 23 HleSeraArgLeuCysser 29
DB 1224 ATCAGTCGCCTGTGTGTCG 1244
RESULT 10
US-09-368-431-13/C
Sequence 13, Application US/09368431
Patent No. 6294651
GENERAL INFORMATION:
APPLICANT: Takashi OKADO et al.
TITLE OF INVENTION: GENE REGULATING AUREOBASIDIN SENSITIVITY
TITLE OF INVENTION: (AS AMENDED)
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: WENDROTH, LIND & PONACK, L.L.P.
STREET: 2033 K Street, N.W., #800
CITY: Washington

STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/368,431
FILING DATE: August 5, 1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/716,873
FILING DATE: September 20, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
TELEX:
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 2856 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ANTI-SENSE: Yes
US-09-368-431-13
Alignment Scores:
Pred. No.: 396
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 5.98%
DB: 4
Gaps: 0
US-09-092-297-17 (1-117) x US-09-368-431-13 (1-2856)
QY 23 HleSeraArgLeuCysser 29
DB 1633 ATCAGTCGCCTGTGTGTCG 1613
RESULT 11
US-09-368-431-14/C
Sequence 14, Application US/09368431
Patent No. 6294651
GENERAL INFORMATION:
APPLICANT: Takashi OKADO et al.
TITLE OF INVENTION: GENE REGULATING AUREOBASIDIN SENSITIVITY
TITLE OF INVENTION: (AS AMENDED)
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: WENDROTH, LIND & PONACK, L.L.P.
STREET: 2033 K Street, N.W., #800
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/368,431
FILING DATE: August 5, 1999
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/716,873
FILING DATE: September 20, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2856 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: mRNA
ANTI-SENSE: Yes
US-09-368-431-14

Alignment Scores:
Pred. No.: 396 Length: 2856
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
Gaps: 0
DB: 4

US-09-092-297-17 (1-117) x US-09-368-431-14 (1-2856)

QY 23 lleserArgLeuLeucysSer 29
Db 1633 ATCAGTCGCTGTGTGTCG 1613

RESULT 12
US-09-041-886-10
Sequence 10, Application US/09041886
Patent No. 6235872
GENERAL INFORMATION:
APPLICANT: Bredesen, Dale E.
APPLICANT: RabiZadeh, SharroZ
TITLE OF INVENTION: Prapoptotic Peptides, Dependence
TITLE OF INVENTION: Polypeptides and Methods of Use
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,886
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 2626
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8449
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 3715 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 532..3286
US-09-041-886-10

Alignment Scores:
Pred. No.: 508 Length: 3715
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
Gaps: 0
DB: 4

US-09-092-297-17 (1-117) x US-09-041-886-10 (1-3715)

QY 43 LeuGysGlnProHisLysArg 49
Db 3620 CTGTCCAGCCACACAAACGT 3640

RESULT 13
US-09-541-782-1/C
Sequence 1, Application US/09541782
Patent No. 6284480
GENERAL INFORMATION:
APPLICANT: Nislow, Corey
APPLICANT: Sakowicz, Roman
APPLICANT: Bernad, Christophe
TITLE OF INVENTION: Antifungal Assay
FILE REFERENCE: 1015
CURRENT APPLICATION NUMBER: US/09/541,782
CURRENT FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 4014
TYPE: DNA
ORGANISM: A.nidulans
US-09-541-782-1

Alignment Scores:
Pred. No.: 546 Length: 4014
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
Gaps: 0
DB: 4

US-09-092-297-17 (1-117) x US-09-541-782-1 (1-4014)

QY 66 AlaValValProLeuAlaArg 72
Db 24 GCTGTGTTCGTCGTGGCCAGA 4

RESULT 14
US-08-570-311-7
Sequence 7, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Progulski-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Leplne, Guylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL

COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE: US/08/570,311
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4080 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 87..3347
US-08-570-311-7
Alignment Scores:
Pred. No.: 554
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 5.98%
DB: 1
Gaps: 0
US-09-092-297-17 (1-117) x US-08-570-311-7 (1-4080)
QY 25 ArgLeuLeuGysSerHisGly 31
|||||
Db 3284 CGGCTACTATGCGATCATGT 3304
RESULT 15
US-08-353-485-7
Sequence 7, Application US/08353485
Patent No. 5830710
GENERAL INFORMATION:
APPLICANT: Proguiske-Tox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepline, Guylatine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville

STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C2
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4080 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 87..3347
US-08-353-485-7
Alignment Scores:
Pred. No.: 554
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 5.98%
DB: 2
Gaps: 0
US-09-092-297-17 (1-117) x US-08-353-485-7 (1-4080)
QY 25 ArgLeuLeuGysSerHisGly 31
|||||
Db 3284 CGGCTACTATGCGATCATGT 3304
Search completed: November 3, 2002, 05:06:19
Job time : 55 secs